

Guía de administración y rendimiento  
de IBM SPSS Modeler Server 14.2



*Nota:* Antes de utilizar esta información y el producto, lea la información general en Avisos el p. .

Este documento contiene información propiedad de SPSS Inc, an IBM Company. Se proporciona con un contrato de licencia y está protegido por leyes de derechos de autor. La información que contiene esta publicación no incluye garantías del producto y cualquier declaración de este manual no se debe considerar como tal.

Al enviar información a IBM o SPSS, el usuario concede a IBM y a SPSS el derecho no exclusivo de utilizar o distribuir la información de la forma que estime adecuada sin incurrir en obligaciones con el usuario.

© **Copyright IBM Corporation 1994, 2011..**

---

# Prefacio

IBM® SPSS® Modeler es el conjunto de programas de minería de datos de IBM Corp. orientado a las empresas. SPSS Modeler ayuda a las organizaciones a mejorar la relación con sus clientes y los ciudadanos a través de la comprensión profunda de los datos. Las organizaciones utilizan la comprensión que les ofrece SPSS Modeler para retener a los clientes más rentables, identificar las oportunidades de venta cruzada, atraer a nuevos clientes, detectar el fraude, reducir el riesgo y mejorar la prestación de servicios del gobierno.

La interfaz visual de SPSS Modeler invita a la pericia empresarial específica de los usuarios, lo que deriva en modelos predictivos más eficaces y la reducción del tiempo necesario para encontrar soluciones. SPSS Modeler ofrece muchas técnicas de modelado tales como pronósticos, clasificaciones, segmentación y algoritmos de detección de asociaciones. Una vez que se crean los modelos, IBM® SPSS® Modeler Solution Publisher permite su distribución en toda la empresa a los encargados de tomar las decisiones o a una base de datos.

## ***Acerca de IBM Business Analytics***

El software IBM Business Analytics ofrece información completa, coherente y precisa en la que los órganos de toma de decisiones confían para mejorar el rendimiento comercial. Un conjunto integral de [inteligencia empresarial](#), [análisis predictivo](#), [rendimiento comercial y gestión de estrategias](#), así como de [aplicaciones de análisis](#) le ofrece una información clara, inmediata e interactiva del rendimiento actual y la capacidad para predecir resultados futuros. En combinación con extensas soluciones sectoriales, prácticas probadas y servicios profesionales, las organizaciones de cualquier tamaño pueden conseguir el máximo de productividad y alcanzar mejores resultados.

Como parte de esta familia, el software de análisis predictivo de IBM SPSS ayuda a las organizaciones a predecir eventos futuros y actuar proactivamente según esa información para lograr mejores resultados comerciales. Los clientes comerciales, gubernamentales y académicos de todo el mundo confían en la tecnología de IBM SPSS como ventaja ante la competencia para atraer, retener y hacer crecer los clientes, reduciendo al mismo tiempo el fraude y reduciendo el riesgo. Al incorporar el software de IBM SPSS en sus operaciones diarias, las organizaciones se convierten en empresas predictivas, capaces de dirigir y automatizar decisiones para alcanzar los objetivos comerciales y lograr una ventaja considerable sobre la competencia. Para obtener más información o contactar con un representante, visite <http://www.ibm.com/spss>.

## ***Asistencia técnica***

La asistencia técnica está disponible para el mantenimiento de los clientes. Los clientes podrán ponerse en contacto con el servicio de asistencia técnica si desean recibir ayuda sobre la utilización de los productos de IBM Corp. o sobre la instalación en los entornos de hardware admitidos. Para ponerse en contacto con el servicio de asistencia, visite el IBM Corp. sitio Web en <http://www.ibm.com/support>. Prepárese para identificarse, identificar a su organización y su acuerdo de asistencia al solicitar asistencia.

---

# Contenido

<b>1</b>	<b><i>Acerca de IBM SPSS Modeler</i></b>	<b>1</b>
	IBM SPSS Modeler Server . . . . .	1
	Opciones de IBM SPSS Modeler . . . . .	2
	IBM SPSS Text Analytics . . . . .	2
	Documentación de IBM SPSS Modeler. . . . .	2
	Ejemplos de aplicaciones . . . . .	4
	Carpeta Demos. . . . .	5
<b>2</b>	<b><i>Arquitectura y recomendaciones de hardware</i></b>	<b>6</b>
	Arquitectura de IBM SPSS Modeler . . . . .	6
	Descripción de Arquitectura. . . . .	6
	Recomendaciones de Hardware. . . . .	8
	Requisitos de RAM y espacio temporal en disco . . . . .	9
	Acceso a los datos . . . . .	11
	Referencia a archivos de datos . . . . .	12
	Importación de archivos de datos de IBM SPSS Statistics . . . . .	13
	Instrucciones para la instalación . . . . .	13
<b>3</b>	<b><i>Asistencia de IBM SPSS Modeler</i></b>	<b>14</b>
	Conexión con IBM SPSS Modeler Server . . . . .	14
	Adición y edición de la conexión de IBM SPSS Modeler Server . . . . .	16
	Búsqueda de servidores en IBM SPSS Collaboration and Deployment Services. . . . .	17
	Sistemas de datos y archivos . . . . .	18
	Autenticación del usuario. . . . .	19
	Creación de archivos. . . . .	19
	Diferencias en los resultados . . . . .	20
<b>4</b>	<b><i>Administración de IBM SPSS Modeler Server</i></b>	<b>21</b>
	Inicio y parada de IBM SPSS Modeler Server. . . . .	21
	Para iniciar, detener y comprobar el estado en Windows. . . . .	21
	Para iniciar, detener y comprobar el estado en UNIX. . . . .	21
	Gestión de procesos del servidor que no responden (sistemas UNIX) . . . . .	22

Administración . . . . .	23
Uso de IBM SPSS Modeler Administration Console . . . . .	24
Inicie Modeler Administration Console . . . . .	24
Configuración de acceso con Modeler Administration Console . . . . .	25
Conexiones de SPSS Modeler Server . . . . .	26
Configuración de SPSS Modeler Server. . . . .	27
Supervisión de SPSS Modeler Server . . . . .	33
Uso del archivo options.cfg. . . . .	34
Cierre de conexiones de base de datos sin utilizar . . . . .	35
Ubicación de licencia de IBM SPSS Statistics . . . . .	36
Uso de SSL para transferencias de datos seguras . . . . .	36
Cómo funciona SSL . . . . .	36
Comunicaciones cliente-servidor y servidor-servidor seguras con SSL . . . . .	37
Configuración de prefijo URL . . . . .	41
Seguridad LDAP con SSL . . . . .	41
Registro del servidor. . . . .	42

## **5 Conceptos básicos de rendimiento 44**

Configuración de rendimiento y optimización del servidor. . . . .	44
Configuración de rendimiento y optimización de cliente . . . . .	44
Optimización y uso de bases de datos. . . . .	47
Optimización de SQL . . . . .	48

## **6 Optimización de SQL 49**

Cómo funciona la generación de SQL . . . . .	50
Ejemplo de generación de SQL . . . . .	51
Configuración de la optimización de SQL . . . . .	52
Presentación preliminar del SQL generado . . . . .	53
Visualización de SQL de nuggets de modelo . . . . .	54
Sugerencias para maximizar la generación de SQL. . . . .	55
Nodos que admiten la generación de SQL. . . . .	56
Operadores y expresiones CLEM que admiten la generación de SQL. . . . .	60
Uso de funciones SQL en expresiones CLEM . . . . .	62
Escritura de consultas SQL. . . . .	62

## **Apéndices**

<b>A</b>	<b>Configuración de Oracle para plataformas UNIX</b>	<b>64</b>
	Configuración de Oracle para la optimización de SQL .....	64
<b>B</b>	<b>Configuración de procesos de inicio de UNIX</b>	<b>66</b>
	Introducción .....	66
	Procesos .....	66
	Inicio y parada automáticos de IBM SPSS Modeler Server .....	67
	Inicio y parada manuales de IBM SPSS Modeler Server .....	67
	Edición de procesos .....	67
	Control de permisos en la creación de archivos .....	68
	IBM SPSS Modeler Server y el paquete Data Access Pack .....	68
	Solución de problemas de la configuración de Connect for ODBC .....	71
	Rutas de bibliotecas .....	74
	Configuración de un controlador para el nodo Enterprise View .....	75
<b>C</b>	<b>Ejecución como un proceso no root en UNIX.</b>	<b>76</b>
	Introducción .....	76
	Configuración de IBM SPSS Modeler Server como un proceso no root. ....	76
<b>D</b>	<b>Equilibrado de cargas con conglomerados de servidores</b>	<b>79</b>
<b>E</b>	<b>Avisos</b>	<b>81</b>
	<b>Índice</b>	<b>84</b>

# ***Acerca de IBM SPSS Modeler***

IBM® SPSS® Modeler es un conjunto de herramientas de minería de datos que permite desarrollar rápidamente modelos predictivos mediante técnicas empresariales y utilizarlos en operaciones empresariales para mejorar la toma de decisiones. Con un diseño que sigue el modelo CRISP-DM, estándar del sector, SPSS Modeler admite el proceso completo de minería de datos, desde los propios datos hasta obtener los mejores resultados empresariales.

SPSS Modeler ofrece una gran variedad de métodos de modelado procedentes del aprendizaje automático, la inteligencia artificial y el estadístico. Los métodos disponibles en la paleta de modelado permiten derivar nueva información procedente de los datos y desarrollar modelos predictivos. Cada método tiene ciertos puntos fuertes y es más adecuado para determinados tipos de problemas.

SPSS Modeler puede adquirirse como producto independiente o utilizarse en conjunto con SPSS Modeler Server. También hay disponible cierto número de opciones adicionales que se resumen en las siguientes secciones. Si desea obtener más información, consulte <http://www.ibm.com/software/analytics/spss/products/modeler/>.

## ***IBM SPSS Modeler Server***

SPSS Modeler utiliza una arquitectura de cliente/servidor para distribuir peticiones de cliente para operaciones que requieren un uso intensivo de los recursos a un software de servidor de gran potencia, lo que proporciona un rendimiento más rápido con conjuntos de datos de mayor volumen. También puede haber disponibles productos o actualizaciones adicionales que no se incluyan en esta lista. Si desea obtener más información, consulte <http://www.ibm.com/software/analytics/spss/products/modeler/>.

**SPSS Modeler.** SPSS Modeler es una versión completamente funcional del producto que se instala y ejecuta en el ordenador de escritorio del usuario. Esta versión se puede ejecutar en modo local como un producto independiente o en modo distribuido junto con IBM® SPSS® Modeler Server para mejorar el rendimiento a la hora de trabajar con grandes conjuntos de datos.

**SPSS Modeler Server.** SPSS Modeler Server se ejecuta ininterrumpidamente en modo de análisis distribuido junto con una o varias instalaciones de IBM® SPSS® Modeler, lo que ofrece un mayor rendimiento cuando se trabaja con grandes conjuntos de datos, ya que las operaciones que requieren un uso intensivo de la memoria se pueden realizar en el servidor sin tener que descargar datos en el equipo cliente. SPSS Modeler Server también ofrece compatibilidad con las capacidades de optimización de SQL y modelado en la base de datos, lo que ofrece ventajas adicionales de rendimiento y automatización. Para ejecutar un análisis debe haber al menos una instalación de SPSS Modeler.

## Opciones de IBM SPSS Modeler

Es posible adquirir una licencia de uso de los siguientes componentes y características que pueden utilizarse con SPSS Modeler. Recuerde que también puede haber disponibles productos o actualizaciones adicionales. Si desea obtener más información, consulte <http://www.ibm.com/software/analytics/spss/products/modeler/>.

- Acceso a SPSS Modeler Server, que ofrece una mayor escalabilidad y rendimiento en conjuntos de datos grandes, así como compatibilidad con las capacidades de optimización de SQL y modelado en la base de datos.
- SPSS Modeler Solution Publisher, permite la puntuación automática o en tiempo real fuera del entorno de SPSS Modeler. Si desea obtener más información, consulte el tema [IBM SPSS Modeler Solution Publisher en el capítulo 2 en IBM SPSS Modeler 14.2 Solution Publisher](#).
- Adaptadores para permitir la distribución en IBM SPSS Collaboration and Deployment Services o la aplicación IBM SPSS Modeler Advantage de baja intensidad. Si desea obtener más información, consulte el tema [Almacenamiento y recuperación de objetos de IBM SPSS Collaboration and Deployment Services Repository en el capítulo 9 en Manual de usuario de IBM SPSS Modeler 14.2](#).

## IBM SPSS Text Analytics

IBM® SPSS® Text Analytics es un complemento totalmente integrado en SPSS Modeler que utiliza tecnologías de lingüística avanzada y NLP para procesar con rapidez una gran variedad de datos de texto sin estructurar, extraer y organizar los conceptos clave y agruparlos en categorías. Las categorías y conceptos extraídos se pueden combinar con los datos estructurados existentes, como pueden ser datos demográficos, y se pueden aplicar para modelar utilizando el conjunto completo de herramientas de minería de datos de IBM® SPSS® Modeler para tomar decisiones mejores y más certeras.

- El nodo Text Mining ofrece modelado de conceptos y categorías así como un programa interactivo donde se puede realizar una exploración avanzada de conglomerados y vínculos de texto, crear sus propias categorías y refinar las plantillas de recursos lingüísticos.
- Hay diversos formatos de importación compatibles, incluyendo blogs y otros orígenes basados en Web.
- También se incluyen plantillas, bibliotecas y diccionarios personalizados para dominios específicos, como puede ser la terminología CRM y genómica.

*Nota:* Es necesario disponer de una licencia independiente para acceder a este componente. Si desea obtener más información, consulte <http://www.ibm.com/software/analytics/spss/products/modeler/>.

## Documentación de IBM SPSS Modeler

Tiene a su disposición una completa documentación en formato de ayuda en línea desde el menú Ayuda de SPSS Modeler. Se incluye documentación para SPSS Modeler, SPSS Modeler Server y SPSS Modeler Solution Publisher, así como el Manual de aplicaciones y otros materiales de apoyo.

La documentación completa de cada producto en formato PDF está disponible en la carpeta *Documentation* en cada DVD del producto.

- **Manual del usuario de IBM SPSS Modeler.** Introducción general sobre cómo usar SPSS Modeler, incluyendo cómo crear rutas de datos, tratar valores perdidos, crear expresiones CLEM, trabajar con proyectos e informes y empaquetar rutas para su distribución en IBM SPSS Collaboration and Deployment Services, Predictive Applications o IBM SPSS Modeler Advantage.
- **Nodos Origen, Proceso y Resultado de IBM SPSS Modeler.** Descripciones de todos los nodos utilizados para leer, procesar y dar salida a datos en diferentes formatos. En la práctica, esto implica todos los nodos que no sean nodos de modelado.
- **Nodos de modelado de IBM SPSS Modeler.** Descripciones de todos los nodos utilizados para crear modelos de minería de datos. IBM® SPSS® Modeler ofrece una variedad de métodos de modelado tomados del aprendizaje de las máquinas, la inteligencia artificial y la estadística. [Si desea obtener más información, consulte el tema Conceptos básicos sobre nodos de modelado en el capítulo 3 en \*Nodos de modelado de IBM SPSS Modeler 14.2\*.](#)
- **Manual de algoritmos de IBM SPSS Modeler.** Descripciones de los fundamentos matemáticos de los métodos de modelado que se utilizan en SPSS Modeler.
- **Manual de aplicaciones de IBM SPSS Modeler.** Los ejemplos de esta guía ofrecen introducciones breves y concisas a métodos y técnicas de modelado específicos. También tiene a su disposición una versión en línea de este manual en el menú Ayuda. [Si desea obtener más información, consulte el tema Ejemplos de aplicaciones en \*Manual de usuario de IBM SPSS Modeler 14.2\*.](#)
- **Procesos y automatización de IBM SPSS Modeler.** Información sobre la automatización del sistema a través de procesos, incluidas las propiedades que se pueden utilizar para manipular nodos y rutas.
- **IBM SPSS Modeler Manual de distribución.** Información sobre la ejecución de rutas y escenarios de SPSS Modeler como pasos en trabajos de procesamiento en IBM® SPSS® Collaboration and Deployment Services Deployment Manager.
- **Guía del desarrollador de IBM SPSS Modeler CLEF.** CLEF permite integrar programas de otros fabricantes, como rutinas de procesamiento de datos o algoritmos de modelado como nodos en SPSS Modeler.
- **Manual de minería interna de bases de datos de IBM SPSS Modeler.** Este manual incluye información sobre cómo utilizar la potencia de su base de datos, tanto para mejorar su rendimiento como para ampliar su oferta de capacidades analíticas a través de algoritmos de terceros.
- **Guía de IBM SPSS Modeler Server y su rendimiento.** Información sobre la configuración y administración de IBM® SPSS® Modeler Server.
- **Manual del usuario de IBM SPSS Modeler Administration Console.** Información sobre cómo instalar y utilizar la interfaz de usuario de la consola para supervisar y configurar SPSS Modeler Server. La consola se implementa como complemento de la aplicación Deployment Manager.

- **Manual de IBM SPSS Modeler Solution Publisher.** SPSS Modeler Solution Publisher es un componente complementario que permite a las organizaciones publicar rutas para su uso fuera del entorno estándar de SPSS Modeler.
- **Manual de CRISP-DM de IBM SPSS Modeler.** Manual que explica paso a paso cómo utilizar la metodología de CRISP-DM en la minería de datos con SPSS Modeler.

## ***Ejemplos de aplicaciones***

Mientras que las herramientas de minería de datos de SPSS Modeler pueden ayudar a resolver una amplia variedad de problemas organizativos y empresariales, los ejemplos de la aplicación ofrecen introducciones breves y adaptadas de técnicas y métodos de modelado específicos. Los conjuntos de datos utilizados aquí son mucho más pequeños que los enormes almacenes de datos gestionados por algunos analizadores de datos, pero los conceptos y métodos implicados deberían ser escalables a las aplicaciones reales.

Para acceder a los ejemplos pulsando Ejemplos de aplicación en el menú Ayuda de SPSS Modeler. Los archivos de datos y rutas de muestra se instalan en la carpeta *Demos* en el directorio de instalación del producto. [Si desea obtener más información, consulte el tema Carpeta Demos en \*Manual de usuario de IBM SPSS Modeler 14.2\*.](#)

**Ejemplos de modelado de base de datos.** Consulte los ejemplos que figuran en el Manual de minería interna de bases de datos de *IBM SPSS Modeler*.

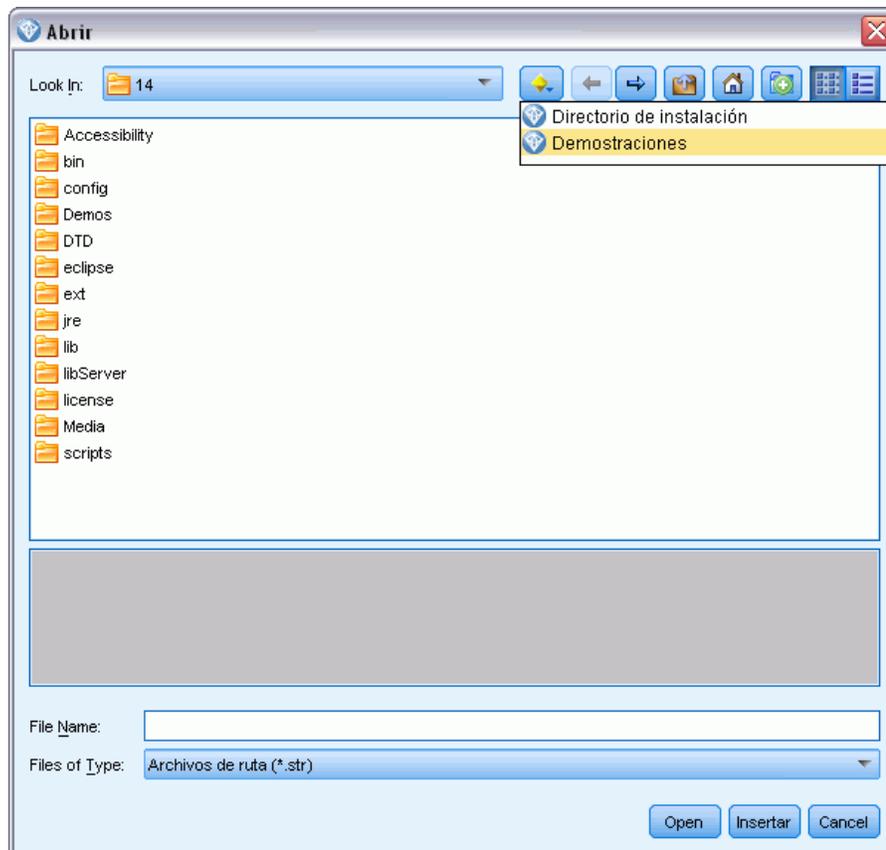
**Ejemplos de procesos.** Consulte los ejemplos que figuran en la Guía de procesos y automatización de *IBM SPSS Modeler*.

## Carpeta Demos

Los archivos de datos y rutas de muestra utilizados con los ejemplos de la aplicación se instalan en la carpeta *Demos* en el directorio de instalación del producto. También puede acceder a esta carpeta desde el grupo de programas IBM SPSS Modeler 14.2 del menú Inicio de Windows o pulsando *Demos* de la lista de directorios recientes en el cuadro de diálogo Abrir archivo.

Figura 1-1

*Selección de la carpeta Demos desde la lista de directorios utilizados recientemente*



# ***Arquitectura y recomendaciones de hardware***

## ***Arquitectura de IBM SPSS Modeler***

Esta sección describe la arquitectura de IBM® SPSS® Modeler Server, incluido el software de servidor, el software cliente y la base de datos. Incluye información sobre cómo está diseñado SPSS Modeler Server para lograr un rendimiento óptimo y proporciona recomendaciones para maximizar este rendimiento mediante la elección de hardware del tamaño apropiado. Concluye con una sección sobre el acceso a datos, que describe dónde instalar los controladores ODBC necesarios.

## ***Descripción de Arquitectura***

IBM® SPSS® Modeler Server utiliza una arquitectura distribuida de tres niveles. Las operaciones de software se comparten entre los equipos cliente y servidor. Son numerosas las ventajas de instalar y utilizar SPSS Modeler Server (frente a IBM® SPSS® Modeler independiente), especialmente al tratar conjuntos de datos grandes:

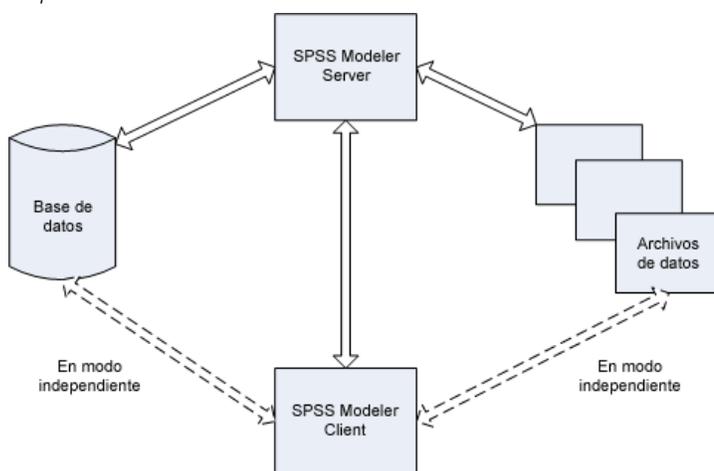
- SPSS Modeler Server se puede ejecutar en UNIX, aparte de Windows, lo que ofrece más flexibilidad al decidir dónde instalarlo. En cualquier plataforma, puede dedicar un equipo servidor más grande y más rápido a los procesos de minería de datos.
- SPSS Modeler Server está optimizado para un rendimiento más rápido. Cuando las operaciones no se pueden incluir en la base de datos, SPSS Modeler Server almacena los resultados intermedios en disco como archivos temporales en lugar de en RAM. Como los servidores suelen tener un considerable espacio disponible en disco, SPSS Modeler Server puede realizar operaciones de ordenación, fusión y agregación en conjuntos muy grandes de datos.
- Utilizando la arquitectura cliente-servidor, puede centralizar los procesos de minería de datos de su organización. La centralización puede ayudar a formalizar el papel de la minería de datos en los procesos empresariales.
- Mediante herramientas de administrador como IBM® SPSS® Collaboration and Deployment Services Deployment Manager (incluida con SPSS Modeler Server) y IBM® SPSS® Collaboration and Deployment Services y IBM® SPSS® Collaboration and Deployment Services Repository (vendidas por separado), puede supervisar los procesos de minería de datos y asegurarse de que estén disponibles recursos de cálculo adecuados. Puede automatizar ciertas tareas de minería de datos, gestionar el acceso a los modelos de datos y compartir los resultados en su organización.

Los componentes de la arquitectura de análisis distribuido de IBM® SPSS® Modeler se muestran en el “Arquitectura de IBM SPSS Modeler Server” gráfico el p. 7.

- **SPSS Modeler.** El software cliente está instalado en el ordenador del usuario final. Proporciona la interfaz de usuario y muestra los resultados de la minería de datos. El cliente es una instalación completa de software de SPSS Modeler, pero cuando se conecta a SPSS Modeler Server para un análisis distribuido, su motor de ejecución no está activo. SPSS Modeler sólo se ejecuta en sistemas operativos Windows.
- **SPSS Modeler Server.** Software de servidor instalado en un equipo servidor, con conectividad de red tanto a SPSS Modeler como a la base de datos. SPSS Modeler Server se ejecuta como un servicio (en Windows) o un proceso demonio (en UNIX), que espera a que los clientes se conecten. Gestiona la ejecución de las rutas y procesos creados con SPSS Modeler.
- **Servidor de base de datos.** El servidor de la base de datos podría ser un almacén de datos activos (por ejemplo, Oracle en un servidor UNIX grande) o, para reducir el impacto sobre otros sistemas operacionales, un mercado de datos en un servidor departamental/local (por ejemplo, SQL Server en Windows).

### Arquitectura de IBM SPSS Modeler Server

Figura 2-1  
Arquitectura de IBM SPSS Modeler Server



Con la arquitectura de análisis distribuido, la mayor parte del procesamiento se realiza en el equipo servidor. Cuando el usuario final ejecuta una ruta, SPSS Modeler envía una descripción de la ruta al servidor. El servidor determina qué operaciones se pueden ejecutar en SQL y crea las consultas adecuadas. Estas consultas se ejecutan en la base de datos y los datos resultantes se pasan al servidor para cualquier procesamiento que no se pueda expresar utilizando SQL. Una vez finalizado el procesamiento, sólo se devuelven al cliente los resultados relevantes.

En caso necesario, SPSS Modeler Server puede ejecutar todas las operaciones de SPSS Modeler fuera de la base de datos. Automáticamente equilibra el uso de la RAM y la memoria en disco para contener datos para su manipulación. Este proceso hace que SPSS Modeler Server sea totalmente compatible con archivos planos.

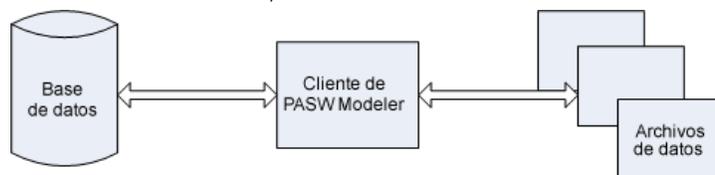
El equilibrado de cargas también está disponible mediante el uso de un conglomerado de servidores para el procesamiento. Los conglomerados están disponibles empezando por IBM SPSS Collaboration and Deployment Services 3.5 a través del complemento Coordinator of Processes. [Si desea obtener más información, consulte el tema Equilibrado de cargas con conglomerados de](#)

servidores en el apéndice D el p. 79. Puede conectarse a un servidor o conglomerado gestionado en Coordinator of Processes directamente a través del cuadro de diálogo Inicio de sesión del servidor de SPSS Modeler. Si desea obtener más información, consulte el tema [Conexión con IBM SPSS Modeler Server](#) en el capítulo 3 en *Manual de usuario de IBM SPSS Modeler 14.2*.

### **Client independiente**

SPSS Modeler también puede configurarse para ejecutarlo como una aplicación local independiente, como muestra el “IBM SPSS Modeler independiente” siguiente gráfico. Si desea obtener más información, consulte el tema [Asistencia de IBM SPSS Modeler](#) en el capítulo 3 el p. 14.

Figura 2-2  
IBM SPSS Modeler independiente



## **Recomendaciones de Hardware**

Al planear la instalación de IBM® SPSS® Modeler Server, se debe considerar el hardware que se utilizará. Aunque SPSS Modeler Server está diseñado para ser rápido, puede maximizar su eficiencia utilizando un hardware del tamaño apropiado para las tareas de minería de datos. Actualizar el hardware es con frecuencia la forma más simple y económica de mejorar el rendimiento general.

**Servidor dedicado.** Instale SPSS Modeler Server en un equipo servidor dedicado donde no compita con otras aplicaciones por los recursos, incluida cualquier base de datos a la que pueda conectarse SPSS Modeler Server. Las operaciones de generación de modelos en concreto requieren un uso intensivo de los recursos y tienen mejor rendimiento cuando no compiten con otras aplicaciones.

*Nota:* Aunque la instalación de SPSS Modeler Server en el mismo ordenador que la base de datos puede reducir el tiempo de transferencia de datos entre la base de datos y el servidor al evitar la transferencia a través de la red, la mayoría de las veces, la mejor configuración es tener la base de datos y el servidor en equipos diferentes para evitar la competición por los recursos. Proporcionar una conexión rápida entre los dos para minimizar el coste de la transferencia de datos.

**Procesadores.** El número de procesadores de la máquina no debe ser menor al número de tareas concurrentes (rutas que se ejecutan simultáneamente) que espera ejecutar de manera habitual. En general, cuantos más procesadores, mejor.

- Una instancia única de SPSS Modeler Server aceptará conexiones con varios clientes (usuarios) y cada conexión de cliente puede iniciar varias ejecuciones de rutas. Por tanto, un servidor puede tener una serie de tareas de ejecución en curso a la vez.
- Como una regla general, utilice un procesador para uno o dos usuarios, dos procesadores para hasta cuatro usuarios y cuatro procesadores para hasta ocho usuarios. A partir de ahí, añada un procesador adicional por cada de dos a cuatro usuarios, según la mezcla de trabajo.

- Si consideramos que parte del procesamiento se puede enviar en la base de datos mediante la optimización de SQL, dos o más usuarios pueden compartir una CPU con una pérdida de rendimiento mínima.
- Las capacidades de ejecutar subprocesos múltiples permiten que varios procesadores ejecuten una única tarea, de manera que añadir procesadores puede mejorar el rendimiento incluso en los casos en que sólo se ejecuta una tarea a la vez. Por lo general, los subprocesos múltiples se utilizan para generar un modelo C5.0 y ciertas operaciones de preparación de datos (ordenar, agregar y fundir).

**plataformas de 64 bits.** Si desea procesar o generar modelos con grandes volúmenes de datos, utilice Solaris, Windows o Linux de 64 bits como plataforma de SPSS Modeler Server y aumente al máximo la cantidad de RAM de la que dispone el ordenador. Para mayores conjuntos de datos, el servidor puede superar rápidamente los límites de memoria por proceso impuestos por plataformas de 32 bits, lo que obliga a volcar datos al disco y aumenta significativamente el tiempo de ejecución. Las implementaciones del servidor de 64 bits pueden beneficiarse de RAM adicional, se recomienda un mínimo de 8 gigabytes (GB). La compatibilidad de 64 bits está disponible para plataformas Solaris, Windows y Linux.

**Necesidades futuras.** Cuando sea viable, asegúrese de que se puede ampliar la memoria y el número de CPU del hardware del servidor, tanto para adaptarse a los aumentos de uso (por ejemplo, mayores números de usuarios simultáneos o aumentos en los requisitos de procesamiento de los usuarios existentes) como para mayores capacidades de subprocesos múltiples de SPSS Modeler Server en el futuro.

## ***Requisitos de RAM y espacio temporal en disco***

IBM® SPSS® Modeler Server utiliza el espacio temporal en disco para procesar grandes volúmenes de datos. La cantidad de espacio temporal que se necesita depende del volumen y el tipo de datos que se procese y del tipo de operaciones que se realicen. El volumen de datos es proporcional al número de filas y de columnas. Cuantas más filas y columnas se procesen, más espacio en disco se necesita.

Esta sección describe las condiciones en las que es necesario el espacio temporal en disco y la RAM adicional, y cómo estimar la cantidad necesaria. Tenga en cuenta que esta sección no trata los requisitos del espacio temporal en disco para los procesos que se realizan en la base de datos, ya que estos requisitos son específicos de cada base de datos.

### ***Condiciones en las que se necesita espacio temporal en disco***

La potente característica de optimización de SQL de IBM® SPSS® Modeler Server permite que el procesamiento se realice en la base de datos (en lugar de en el servidor) siempre que sea posible. Sin embargo, si se cumple alguna de las siguientes condiciones, no se puede utilizar la optimización de SQL:

- Los datos que se van a procesar están almacenados en un archivo plano en lugar de en una base de datos.
- La optimización de SQL está desactivada.
- La operación de procesamiento no se puede optimizar utilizando SQL.

Cuando no se puede utilizar la optimización de SQL, los siguientes nodos de manipulación de datos y funciones CLEM crean copias temporales en disco de todos o parte de los datos. Si las rutas que se utilizan en su empresa contienen estos comandos de procesamiento o funciones, puede necesitar establecer aparte espacio adicional en disco en su servidor.

- nodo Agregar
- nodo Distinguir
- nodo Intervalos
- Nodo Fundir al utilizar la opción de fusión por clave
- Cualquier nodo de modelado
- nodo Ordenar
- Nodo de resultado Tabla
- Funciones @OFFSET en donde la condición de referencia utiliza @THIS.
- Cualquier función @ , como @MIN, @MAX y @AVE, en la que se calcula el parámetro de desplazamiento.

### ***Cálculo de la cantidad de espacio temporal en disco***

En general, IBM® SPSS® Modeler Server necesita poder escribir un archivo temporal que tenga al menos *tres veces el tamaño* del conjunto de datos original. Por ejemplo, si el archivo de datos es de 2 GB y no se utiliza la generación de SQL, SPSS Modeler Server necesitará 6 GB de espacio en disco para procesar los datos. Como cada cuenta de usuario simultánea crea sus propios archivos temporales, necesitará aumentar el espacio en disco en función de cada usuario simultáneo.

Si su empresa utiliza frecuentemente archivos temporales grandes, considere la posibilidad de utilizar un sistema de archivos independiente para los archivos temporales de IBM® SPSS® Modeler, creado en un disco diferente. Si desea obtener mejores resultados, se puede utilizar un RAID 0 o un conjunto de datos en varios volúmenes que abarque varios discos físicos para acelerar las operaciones en disco, idealmente situado cada disco del sistema de archivos en varios volúmenes en un controlador de disco diferente.

### ***Requisitos de RAM***

Para la mayor parte del procesamiento que no se puede realizar en la base de datos, IBM® SPSS® Modeler Server almacena los resultados intermedios en disco como archivos temporales en vez de en la memoria (RAM). Sin embargo, para los nodos de modelado, se utiliza la RAM si es posible. Los nodos Red neuronal, Kohonen y K-medias requieren grandes cantidades de RAM. Si estos nodos se utilizan frecuentemente en su empresa, considere instalar más RAM en el servidor.

En general, el número de bytes de RAM necesarios se puede estimar por

(número\_de\_registros \* número\_de\_casillas\_por\_registro) \* número\_de\_bytes\_por\_casilla

donde number\_of\_cells\_per\_record puede alcanzar un valor muy grande si hay campos nominales.

Consulte la sección sobre los requisitos del sistema del manual de instalación del servidor para conocer las recomendaciones de RAM actuales. Para cuatro usuarios simultáneos o más, se recomienda más RAM. La memoria se debe compartir entre tareas simultáneas, por lo que se debe aumentar en consecuencia. En general, añadir memoria es probablemente una de las maneras más rentables de mejorar el rendimiento general.

## Acceso a los datos

To read or write to a database, you must have an ODBC data source installed and configured for the relevant database, with read or write permissions as needed. The IBM® SPSS® Data Access Pack includes a set of ODBC drivers that can be used for this purpose, and these drivers are available on the IBM SPSS Data Access Pack installation disk shipped with this release. If you have questions about creating or setting permissions for ODBC data sources, contact your database administrator.

Database support in IBM® SPSS® Modeler is classified into three tiers, each representing a different level of support for SQL pushback and optimization, depending on the database vendor. The different levels of support are implemented by means of a number of system settings, which can be customized as part of a Services engagement.

The three tiers of database support are:

Tabla 2-1

*Database support tiers*

Support tier	Description
Nivel 1	Toda la retrotracción SQL posible está disponible, con optimización específica para bases de datos SQL.
Nivel 2	La mayoría de la retrotracción SQL está disponible, con optimización no específica para bases de datos SQL.
Nivel 3	No hay retrotracción u optimización para SQL, sólo la lectura y la escritura de datos en la base de datos.

### Controladores ODBC compatibles

Si desea obtener la información más reciente acerca de las bases de datos y los controladores ODBC compatibles y cuyo uso se ha comprobado con SPSS Modeler 14.2, vea las matrices de compatibilidad de productos disponibles en el sitio de asistencia corporativo (<http://www.ibm.com/support>).

### Dónde instalar los controladores

Tenga en cuenta que todos los controladores ODBC deben instalarse y configurarse en todos los ordenadores en los que pueda realizarse el procesamiento.

- Si desea ejecutar IBM® SPSS® Modeler en modo local (independiente), deberá instalar los controladores en el ordenador local.
- Si ejecuta SPSS Modeler en modo distribuido en un servidor IBM® SPSS® Modeler Server, remoto, los controladores ODBC se deberán instalar en el ordenador en el que se encuentre SPSS Modeler Server.

- Si necesita poder acceder a los mismos orígenes de datos desde SPSS Modeler y SPSS Modeler Server, deberá instalar los controladores ODBC en ambos ordenadores.
- Si desea ejecutar SPSS Modeler a través de Terminal Services, deberá instalar los controladores de ODBC en el servidor de Terminal Services en el que se haya instalado SPSS Modeler.
- Si desea utilizar el tiempo de ejecución de IBM® SPSS® Modeler Solution Publisher para ejecutar las rutas publicadas en otro ordenador, también deberá instalar y configurar los controladores ODBC en dicho ordenador.

**Nota:** Si utiliza SPSS Modeler Server en UNIX para acceder a una base de datos Teradata, debe utilizar ODBC Driver Manager instalado con el controlador Teradata ODBC. Para implementar este cambio en SPSS Modeler Server indique el valor de ODBC\_DRIVER\_MANAGER\_PATH junto al principio del proceso modelersrv.sh, donde indican los comentarios. Esta variable de entorno se debe configurar con la ubicación de ODBC Driver Manager que se incluye junto con el controlador Teradata ODBC (/usr/odbc/lib en una instalación normal del controlador Teradata ODBC). Deberá reiniciar SPSS Modeler Server para que los cambios surtan efecto. Para obtener más información acerca de las plataformas SPSS Modeler Server compatibles en el acceso de Teradata y la versión del controlador ODBC de Teradata compatible, consulte el sitio de asistencia corporativo <http://www.ibm.com/support>.

*Nota:* las reglas anteriores se aplican específicamente al acceso de datos en una base de datos. Otros tipos de operaciones de archivos, como abrir y guardar rutas, proyectos, modelos, nodos, PMML, resultados y archivos de procesos, se realizan siempre en el cliente y siempre se especifican en términos del sistema de archivos del equipo cliente. Además, el comando Definir directorio en SPSS Modeler define el directorio de trabajo para los objetos *locales* del cliente (por ejemplo, rutas), pero no afecta al directorio de trabajo de servidor.

**UNIX.** Si desea obtener más información sobre cómo configurar SPSS Modeler Server en UNIX para trabajar con la tecnología de acceso a datos de IBM® SPSS® Statistics, consulte [Configuración de procesos de inicio de UNIX](#) el p. 66.

## **Referencia a archivos de datos**

**Windows.** Si almacena datos en el mismo ordenador que IBM® SPSS® Modeler Server, le recomendamos que proporcione la ruta a los datos desde la perspectiva del equipo servidor (por ejemplo, *C:\DatosServidor\Ventas 1998.csv*). El rendimiento es más rápido cuando no se utiliza la red para localizar el archivo.

Si los datos se almacenan en un host diferente, le recomendamos utilizar las referencias de archivos UNC (por ejemplo, *\\miservidordatos\DatosServidor\Ventas 1998.csv*). Tenga en cuenta que los nombres UNC sólo funcionan cuando la ruta contiene el nombre de un recurso de red compartido. El ordenador en el que se realiza la referencia debe tener permiso para leer el archivo especificado. Si frecuentemente cambia del modo de análisis local a distribuido, utilice referencias de archivos UNC, ya que funcionan independientemente del modo.

**UNIX.** Para hacer referencia a los archivos de datos que se encuentran en un servidor UNIX, utilice la especificación del archivo completa y las barras diagonales (por ejemplo, */public/datos/DatosServidor/Ventas 1998.csv*). Evite utilizar el carácter de barra invertida en el directorio UNIX y en los nombres de archivos de los datos utilizados con SPSS Modeler

Server. No importa si un archivo de texto utiliza formato UNIX o DOS; ambos se gestionan automáticamente.

### ***Importación de archivos de datos de IBM SPSS Statistics***

Si también está ejecutando IBM® SPSS® Statistics Server en su empresa, es posible que los usuarios quieran importar o exportar datos de SPSS Statistics mientras están en el modo distribuido. Recuerde que cuando IBM® SPSS® Modeler se ejecuta en modo distribuido, presenta el sistema de archivos del servidor. El cliente de SPSS Statistics funciona de la misma manera. Para que se pueda realizar la importación y exportación entre las dos aplicaciones, ambos clientes deben estar trabajando en el mismo modo. Si no lo están, sus vistas del sistema de archivos serán diferentes y no podrán compartir archivos. Los nodos de SPSS Statistics de IBM® SPSS® Modeler pueden iniciar el cliente de SPSS Statistics automáticamente, pero los usuarios deben asegurarse primero de que el cliente de SPSS Statistics está trabajando en el mismo modo que SPSS Modeler.

### ***Instrucciones para la instalación***

Si desea obtener información sobre la instalación de IBM® SPSS® Modeler Server, consulte las instrucciones que encontrará en la carpeta *documentation\installation* del CD de IBM SPSS Modeler Server. Existen documentos diferentes para Windows y UNIX.

Si desea obtener información completa sobre la instalación y uso de IBM® SPSS® Modeler, consulte el CD de IBM SPSS Modeler Server.

# ***Asistencia de IBM SPSS Modeler***

Esta sección está destinada a administradores y personal de soporte técnico que proporcionan asistencia a usuarios de IBM® SPSS® Modeler. Abarca los siguientes temas:

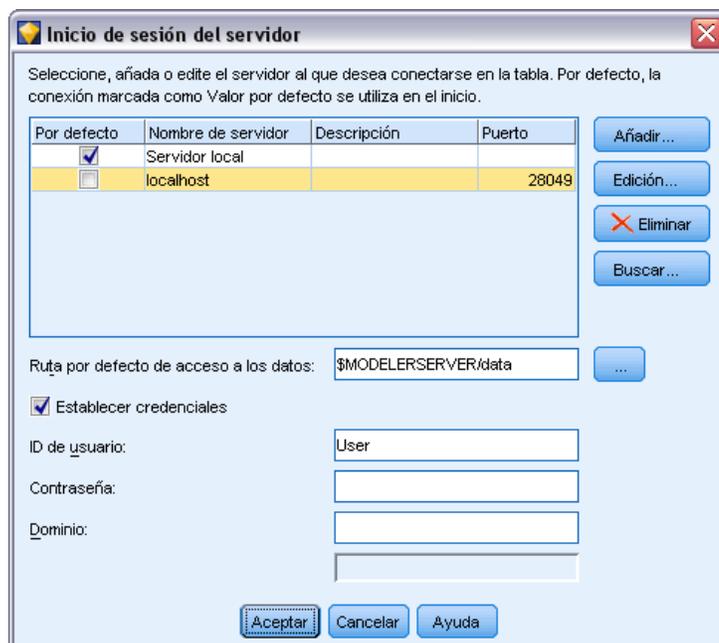
- Cómo iniciar sesión en IBM® SPSS® Modeler Server (o trabajar de manera independiente desconectándose de un servidor)
- Sistemas de datos y archivos que puede que necesiten los usuarios
- Cuentas de usuario y permisos de archivo pertenecientes a SPSS Modeler Server
- Diferencias en el resultado que los usuarios percibirán al alternar entre SPSS Modeler Server y IBM® SPSS® Modeler

## ***Conexión con IBM SPSS Modeler Server***

IBM® SPSS® Modeler puede ejecutarse como una aplicación independiente o como un cliente conectado a IBM® SPSS® Modeler Server directamente o a SPSS Modeler Server o un conglomerado de servidores a través del complemento Coordinator of Processes de IBM® SPSS® Collaboration and Deployment Services. El estado de la conexión actual se muestra en la parte inferior izquierda de la ventana de SPSS Modeler.

Siempre que desee conectarse a un servidor, puede introducir manualmente el nombre de servidor al que desee conectarse o seleccione un nombre que haya definido anteriormente. Sin embargo, si tiene IBM SPSS Collaboration and Deployment Services, puede buscar en una lista de servidores o conglomerados de servidores del cuadro de diálogo Inicio de sesión del servidor. La capacidad de buscar entre los servicios de Statistics que se ejecutan en una red está disponible a través de Coordinator of Processes. [Si desea obtener más información, consulte el tema Equilibrado de cargas con conglomerados de servidores en el apéndice D el p. 79.](#)

Figura 3-1  
Cuadro de diálogo Acceso al servidor



### Para conectar con un servidor

- ▶ En el menú Herramientas, pulse en Inicio de sesión del servidor. Se abre el cuadro de diálogo Inicio de sesión del servidor. Si lo prefiere, pulse dos veces con el ratón en el área de estado de la conexión de la ventana de SPSS Modeler.
- ▶ En el cuadro de diálogo, especifique las opciones para conectarse al equipo servidor local o seleccione una conexión de la tabla.
  - Pulse en Añadir o Edición para añadir o editar una conexión. [Si desea obtener más información, consulte el tema Adición y edición de la conexión de IBM SPSS Modeler Server en Manual de usuario de IBM SPSS Modeler 14.2.](#)
  - Pulse en Buscar para acceder a un servidor o conglomerado de servidores en Coordinator of Processes. [Si desea obtener más información, consulte el tema Búsqueda de servidores en IBM SPSS Collaboration and Deployment Services en Manual de usuario de IBM SPSS Modeler 14.2.](#)

**Tabla Servidor.** Esta tabla contiene el conjunto de conexiones de servidor definidas. La tabla muestra la conexión por defecto, el nombre de servidor, la descripción y el número de puerto. Puede añadir manualmente una nueva conexión, así como seleccionar o buscar una conexión existente. Para establecer un servidor específico como la conexión por defecto, seleccione la casilla de verificación en la columna Por defecto de la tabla para la conexión.

**Ruta por defecto de acceso a los datos.** Especifique la ruta utilizada para los datos del equipo servidor. Pulse en el botón de puntos suspensivos (...) para examinar la ubicación deseada.

**Establecer credenciales.** Deje esta casilla sin seleccionar para activar la función de **inicio de sesión único**, que tratará de iniciar la sesión del usuario en el servidor con los detalles de nombre de usuario y contraseña del equipo local. Si no es posible el inicio único de sesión o si selecciona esta casilla para desactivar el inicio único de sesión (por ejemplo, para iniciar la sesión en una cuenta de administrador), tendrá activados los siguientes campos para que introduzca las credenciales.

**ID de usuario.** Introduzca el nombre de usuario con el que se inicia sesión en el servidor.

**Contraseña.** Introduzca la contraseña asociada al nombre de usuario especificado.

**Dominio.** Especifique el dominio utilizado para iniciar sesión en el servidor. El nombre de dominio es obligatorio sólo si el equipo servidor está en un dominio de Windows distinto que el del equipo cliente.

- ▶ Pulse en Aceptar para completar la conexión.

### **Desconexión de un servidor**

- ▶ En el menú Herramientas, pulse en Inicio de sesión del servidor. Se abre el cuadro de diálogo Inicio de sesión del servidor. Si lo prefiere, pulse dos veces con el ratón en el área de estado de la conexión de la ventana de SPSS Modeler.
- ▶ En el cuadro de diálogo, seleccione el Servidor local y pulse en Aceptar.

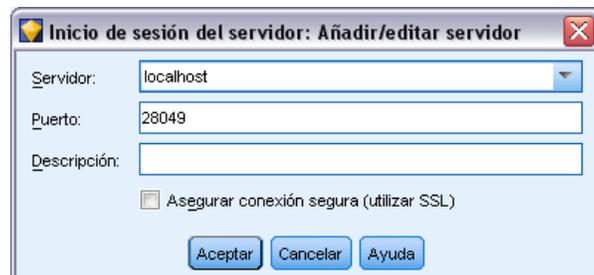
## **Adición y edición de la conexión de IBM SPSS Modeler Server**

Puede editar o añadir manualmente una conexión de servidor en el cuadro de diálogo Inicio de sesión del servidor. Si pulsa en Añadir, puede acceder al cuadro de diálogo Añadir/editar servidor vacío en el que puede introducir los detalles de conexión de servidor. Al seleccionar una conexión existente y pulsar en Editar en el cuadro de diálogo Inicio de sesión del servidor, se abre el cuadro de diálogo Añadir/editar servidor con los detalles de dicha conexión de modo que puede realizar cualquier cambio.

*Nota:* No puede editar una conexión de servidor que se haya añadido desde IBM® SPSS® Collaboration and Deployment Services, ya que el nombre, puerto y otros detalles se definen en IBM SPSS Collaboration and Deployment Services.

**Figura 3-2**

*Cuadro de diálogo Inicio de sesión del servidor: Añadir/editar servidor*



### **Adición de conexiones de servidor**

- ▶ En el menú Herramientas, pulse en Inicio de sesión del servidor. Se abre el cuadro de diálogo Inicio de sesión del servidor.
- ▶ En este cuadro de diálogo, pulse en Añadir. Se abre el cuadro de diálogo Inicio de sesión del servidor: Añadir/editar servidor.
- ▶ Introduzca los detalles de conexión de servidor y pulse en Aceptar para guardar la conexión y volver al cuadro de diálogo Inicio de sesión del servidor.
  - **Servidor.** Especifique un servidor disponible o seleccione uno de la lista. El equipo servidor se puede identificar por un nombre alfanumérico (por ejemplo, *miservidor*) o por una dirección IP asignada al equipo servidor (por ejemplo, 202.123.456.78).
  - **Puerto.** Especifique el número de puerto en el que el servidor escucha. Si no funciona el número de puerto por defecto, solicite el número de puerto correcto al administrador del sistema.
  - **Descripción.** Introduzca una descripción opcional para esta conexión de servidor.
  - **Asegurar conexión segura (utilizar SSL).** Especifica si se debe usar una conexión SSL (del inglés **Secure Sockets Layer**, capa de sockets seguros). SSL es un protocolo normalmente utilizado para asegurar el conjunto de datos que se envía a través de una red. Para utilizar esta función, SSL debe estar activado en el servidor que aloja IBM® SPSS® Modeler Server. Si es preciso, póngase en contacto con el administrador local para obtener más detalles.

### **Edición de conexiones de servidor**

- ▶ En el menú Herramientas, pulse en Inicio de sesión del servidor. Se abre el cuadro de diálogo Inicio de sesión del servidor.
- ▶ En este cuadro de diálogo, seleccione la conexión que desee editar y, a continuación, pulse en Editar. Se abre el cuadro de diálogo Inicio de sesión del servidor: Añadir/editar servidor.
- ▶ Cambie los detalles de conexión de servidor y pulse en Aceptar para guardar los cambios y volver al cuadro de diálogo Inicio de sesión del servidor.

## **Búsqueda de servidores en IBM SPSS Collaboration and Deployment Services**

En lugar de introducir una conexión de servidor manualmente, puede seleccionar un servidor o conglomerado de servidores disponible en la red a través de Coordinator of Processes, disponible en IBM® SPSS® Collaboration and Deployment Services. Un conglomerado de servidores es un grupo de servidores entre los que Coordinator of Processes determina el servidor más adecuado para responder a una solicitud de procesamiento. [Si desea obtener más información, consulte el tema Equilibrado de cargas con conglomerados de servidores en el apéndice D el p. 79.](#)

Aunque puede añadir servidores manualmente al cuadro de diálogo Inicio de sesión del servidor, la búsqueda de servidores disponibles le permite conectarse a servidores sin que sea necesario que conozca el nombre de servidor y número de puerto correctos. Esta información se proporciona automáticamente. Sin embargo, todavía necesita la información de inicio de sesión correcta, como el nombre de usuario, dominio y contraseña.

*Nota:* Si no tiene acceso a la capacidad Coordinator of Processes, todavía puede introducir manualmente el nombre de servidor al que desee conectarse o seleccionar un nombre que haya definido anteriormente. [Si desea obtener más información, consulte el tema Adición y edición de la conexión de IBM SPSS Modeler Server en Manual de usuario de IBM SPSS Modeler 14.2.](#)

Figura 3-3  
Cuadro de diálogo Buscar servidores



### **Búsqueda de servidores y conglomerados**

- ▶ En el menú Herramientas, pulse en Inicio de sesión del servidor. Se abre el cuadro de diálogo Inicio de sesión del servidor.
- ▶ En este cuadro de diálogo, pulse en Buscar para abrir el cuadro de diálogo Buscar servidores. Si no ha iniciado sesión en IBM SPSS Collaboration and Deployment Services cuando intente buscar en Coordinator of Processes, se le pedirá que lo haga. [Si desea obtener más información, consulte el tema Conexión con IBM SPSS Collaboration and Deployment Services Repository en el capítulo 9 en Manual de usuario de IBM SPSS Modeler 14.2.](#)
- ▶ Seleccione el servidor o el conglomerado de servidores de la lista.
- ▶ Pulse en Aceptar para cerrar el cuadro de diálogo y añadir esta conexión a la tabla en el cuadro de diálogo Inicio de sesión del servidor.

## **Sistemas de datos y archivos**

Puede que aquellos usuarios que trabajen con IBM® SPSS® Modeler Server necesiten acceder a archivos de datos y otros orígenes de datos de la red, así como guardar archivos en la red. Por lo tanto, es posible que necesiten la siguiente información cuando proceda:

- **Información sobre los orígenes de datos ODBC.** En caso de que sea necesario que los usuarios accedan a los orígenes de datos ODBC definidos en el equipo servidor, necesitarán los nombres, descripciones e información de inicio de sesión (incluidos los ID y contraseñas de inicio de sesión en la base de datos) relativa a los orígenes de datos.

- **Acceso a archivos de datos.** En caso de que los usuarios precisen acceder a archivos de datos que se encuentren en el equipo servidor o en cualquier otra parte de la red, necesitarán los nombres y ubicaciones de los archivos de datos.
- **Ubicación de archivos guardados.** Cuando un usuario guarda datos mientras está conectado a SPSS Modeler Server, estará tratando de guardar archivos en el equipo servidor. No obstante, éste suele estar protegido contra escritura. Si así es, indique dónde deberían guardarse los archivos de datos. (Generalmente, la ubicación es el directorio principal del usuario.)

## Autenticación del usuario

IBM® SPSS® Modeler Server utiliza el sistema operativo del equipo servidor para autenticar a aquellos usuarios que se conectan al servidor. Cuando un usuario se conecta a SPSS Modeler Server, todas las operaciones que se llevan a cabo en nombre del usuario tendrán lugar dentro del contexto de seguridad del usuario. El acceso a las tablas de una base de datos estará sujeto a los privilegios de usuario y/o contraseña de la base de datos en sí.

**Windows.** En Windows, cualquier usuario con una cuenta válida en la red de host puede iniciar sesión. Con la autenticación por defecto, los usuarios deben tener derechos en el directorio `<modeler_server_install>\Tmp`. Sin estos derechos, los usuarios no pueden iniciar sesión en SPSS Modeler Server desde el cliente utilizando la autenticación por defecto de Windows.

**UNIX.** Por defecto, se da por hecho que SPSS Modeler Server se ejecuta como root en UNIX, lo que hace posible que cualquier usuario con una cuenta válida en la red del host inicie sesión, si bien limita el acceso de dicho usuario a sus propios archivos y directorios. No obstante, puede configurar SPSS Modeler Server para que se ejecute sin privilegios de root. En tal caso, deberá crear una base de datos privada de contraseñas para usarla con fines de autenticación y todos los usuarios de IBM® SPSS® Modeler compartirán una única cuenta de usuario de UNIX (y, en consecuencia, el acceso a los archivos de datos). [Si desea obtener más información, consulte el tema Configuración de IBM SPSS Modeler Server como un proceso no root. en el apéndice C el p. 76.](#)

En plataformas Solaris, HP-UX, Linux y AIX, SPSS Modeler Server utiliza PAM para la autenticación. Puede usar el nombre de servicio `modelerserver` para configurar los módulos PAM para SPSS Modeler Server en caso necesario.

## Creación de archivos

Cuando IBM® SPSS® Modeler Server accede a datos y los procesa, a menudo tiene que guardar una copia temporal de dichos datos en disco. La cantidad de espacio en disco que se va a usar para los archivos temporales dependerá del tamaño del archivo de datos que el usuario final está analizando y del tipo de análisis realizado. [Si desea obtener más información, consulte el tema Requisitos de RAM y espacio temporal en disco en el capítulo 2 el p. 9.](#)

**UNIX.** Las versiones de UNIX de SPSS Modeler Server utilizan el comando `umask` de UNIX para establecer los permisos de archivo para los archivos temporales. Puede omitir los permisos por defecto del servidor. [Si desea obtener más información, consulte el tema Control de permisos en la creación de archivos en el apéndice B el p. 68.](#)

## ***Diferencias en los resultados***

Los usuarios que realizan análisis en ambos modos posiblemente perciban una ligera diferencia en los resultados de IBM® SPSS® Modeler y IBM® SPSS® Modeler Server. Esta disparidad suele deberse al orden de los registros o a diferencias en el redondeo.

**Orden de registros.** A menos que una ruta ordene explícitamente los registros, el orden en que los registros se presentan puede diferir entre las rutas ejecutadas localmente y las que se ejecutan en el servidor. Asimismo, puede que surjan diferencias en el orden entre operaciones realizadas en una base de datos y aquellas realizadas en SPSS Modeler Server. Estas diferencias se deben a que cada sistema utiliza algoritmos distintos para implementar las funciones que pueden ordenar los registros, como es el caso de la agregación. Del mismo modo, tenga en cuenta que SQL no especifica el orden en que se obtienen los registros de una base de datos si no existe una operación de ordenación específica.

**Diferencias de redondeo.** Cuando IBM® SPSS® Modeler funciona en modo local, utiliza un formato interno para almacenar valores de coma flotante distinto del de SPSS Modeler Server. Por motivos de diferencias en el redondeo, los resultados pueden ser ligeramente distintos en cada versión.

# Administración de IBM SPSS Modeler Server

Este capítulo contiene información acerca del inicio y parada de IBM® SPSS® Modeler Server, la configuración de diversas opciones de servidor y la interpretación del archivo de registro. Asimismo, explica el modo de usar IBM® SPSS® Collaboration and Deployment Services Deployment Manager, una aplicación que facilita la configuración y supervisión del servidor. Si desea ver las instrucciones de instalación de este componente, consulte las instrucciones de instalación de SPSS Modeler Server que se incluyen en el CD del producto.

## ***Inicio y parada de IBM SPSS Modeler Server***

IBM® SPSS® Modeler Server se ejecuta como un servicio de Windows o como un proceso demonio de UNIX.

*Nota sobre programación:* Al detener SPSS Modeler Server, se desconecta a los usuarios finales y se finalizan las sesiones, de modo que trate de programar los reinicios del servidor en períodos de poco uso. Si esto no es factible, asegúrese de que informa a los usuarios antes de detener el servidor.

### ***Para iniciar, detener y comprobar el estado en Windows***

En Windows, IBM® SPSS® Modeler Server se controla en el cuadro de diálogo Servicios del Panel de control.

- ▶ **Windows XP.** Abra el menú Inicio de Windows. Seleccione Configuración y, a continuación, Panel de control. Pulse dos veces en Herramientas administrativas y en Servicios.

**Windows 2003 ó 2008.** Abra el menú Inicio de Windows. Seleccione Panel de control, Herramientas administrativas y, a continuación, Servicios.

- ▶ Seleccione el servicio de IBM SPSS Modeler Server <nn.n>. Ahora puede comprobar su estado, iniciarlo o detenerlo y editar los parámetros de arranque, según sea adecuado.

Por defecto, el servicio está configurado para iniciarse automáticamente, lo que quiere decir que, si se detiene, se reiniciará automáticamente al reiniciar el ordenador. Cuando se inicia de esta forma, el servicio se ejecuta desatendido y el equipo servidor podrá desconectarse sin que esto afecte en modo alguno.

### ***Para iniciar, detener y comprobar el estado en UNIX***

En UNIX, IBM® SPSS® Modeler Server puede iniciarse y detenerse ejecutando el proceso `modelsrv.sh` en el directorio de instalación de SPSS Modeler Server.

- Cambie al directorio de instalación de SPSS Modeler Server. Por ejemplo, en un símbolo del sistema de UNIX, escriba

```
cd /usr/modelersrv
```

donde *modelersrv* es el directorio de instalación de SPSS Modeler Server.

- Para iniciar el servidor, en el símbolo del sistema, escriba

```
./modelersrv.sh start
```

- Para detener el servidor, en el símbolo del sistema, escriba

```
./modelersrv.sh stop
```

- Para comprobar el estado de SPSS Modeler Server, en el símbolo del sistema de UNIX, escriba

```
./modelersrv.sh list
```

y observe el resultado, que es similar al del comando `ps` de UNIX. El primer proceso de la lista es el proceso demonio de SPSS Modeler Server y el resto, sesiones de IBM® SPSS® Modeler.

El programa de instalación de SPSS Modeler Server incluye un proceso (*auto.sh*) mediante el que se configura el sistema para iniciar automáticamente el proceso demonio del servidor durante el arranque. Si ha ejecutado el proceso y, a continuación, detiene el servidor, el proceso demonio del servidor se reiniciará automáticamente al reiniciar el ordenador. [Si desea obtener más información, consulte el tema Inicio y parada automáticos de IBM SPSS Modeler Server en el apéndice B el p. 67.](#)

### ***límites de kernel de UNIX***

Debe asegurarse de que los límites de kernel del sistema son suficientes para el funcionamiento de SPSS Modeler Server. Los límites u de datos, memoria y archivos son especialmente importantes y deben establecerse como ilimitados dentro del entorno de SPSS Modeler Server. Para ello:

- Añada los siguientes comandos a *modelersrv.sh*:

```
ulimit -d unlimited
```

```
ulimit -m unlimited
```

```
ulimit -f unlimited
```

- Reinicie SPSS Modeler Server.

## ***Gestión de procesos del servidor que no responden (sistemas UNIX)***

Los procesos de IBM® SPSS® Modeler Server pueden dejar de responder por diversas razones, incluidas situaciones en las que realizan una llamada al sistema o una llamada a un controlador ODBC que se bloquea (la llamada nunca se devuelve y tarda mucho en hacerlo). Si los procesos UNIX entran en este estado, se pueden limpiar mediante el comando `kill` de UNIX (las interrupciones iniciadas por el cliente de IBM® SPSS® Modeler o el cierre del cliente de SPSS Modeler no surtirán efecto). Se proporciona un comando `kill` como alternativa al comando `stop`

habitual y permite que un administrador use *modelersrv.sh* para enviar fácilmente el comando kill adecuado.

En sistemas susceptibles de acumular procesos del servidor inutilizables (“inertes”), recomendamos detener SPSS Modeler Server y reiniciarlo a intervalos regulares mediante la siguiente secuencia de comandos:

```
cd directorio_instalación_modeler_server
./modelersrv.sh stop
./modelersrv.sh kill
```

Esos procesos de SPSS Modeler finalizados mediante el comando *modelersrv.sh kill* dejarán archivos temporales (del directorio temporal) que habrá que eliminar manualmente. También pueden quedar archivos temporales en algunas otras situaciones, que incluyen los bloqueos de aplicaciones debidos al agotamiento de los recursos, interrupciones del usuario, bloqueos del sistema u otras razones. Por lo tanto, recomendamos que, como parte del proceso de reiniciar SPSS Modeler Server a intervalos regulares, se eliminen todos los archivos restantes del directorio temporal de SPSS Modeler.

Una vez que se han cerrado todos los procesos del servidor y se han eliminado todos los archivos temporales, se puede reiniciar SPSS Modeler Server con seguridad.

## Administración

IBM® SPSS® Modeler Server cuenta con una serie de opciones configurables con las que se controla su comportamiento. Puede establecer estas opciones de dos formas:

- Utilice la aplicación IBM® SPSS® Collaboration and Deployment Services Deployment Manager, que está disponible de forma gratuita para los clientes actuales de IBM® SPSS® Modeler. [Si desea obtener más información, consulte el tema Uso de IBM SPSS Modeler Administration Console el p. 24.](#)
- Utilice el archivo de texto *options.cfg*, que se encuentra en el directorio *[ruta de instalación del servidor]/config*. [Si desea obtener más información, consulte el tema Uso del archivo options.cfg el p. 34.](#)

Es recomendable que instale y utilice Deployment Manager como herramienta de administración en lugar de editar el archivo *options.cfg* ya que para editar el archivo es necesario acceder al sistema de archivos de SPSS Modeler Server, mientras que Deployment Manager permite autorizar a cualquier persona con una cuenta de usuario para que ajuste estas opciones. Por otra parte, Deployment Manager ofrece información adicional sobre los procesos del servidor, lo que permite supervisar el uso y el rendimiento. Además, y al contrario de lo que sucede al editar el archivo de configuración, la mayoría de las opciones de configuración se pueden cambiar sin tener que reiniciar SPSS Modeler Server.

Encontrará más información sobre el uso de Deployment Manager y el archivo *options.cfg* en las siguientes secciones.

## **Uso de IBM SPSS Modeler Administration Console**

La aplicación Modeler Administration Console proporciona una interfaz de usuario de la consola para supervisar y configurar las instalaciones de SPSS Modeler Server y está disponible de forma completamente gratuita para los clientes actuales de SPSS Modeler Server. La aplicación sólo se puede instalar en los ordenadores con Windows; sin embargo, puede administrar un servidor que esté instalado en cualquier plataforma compatible.

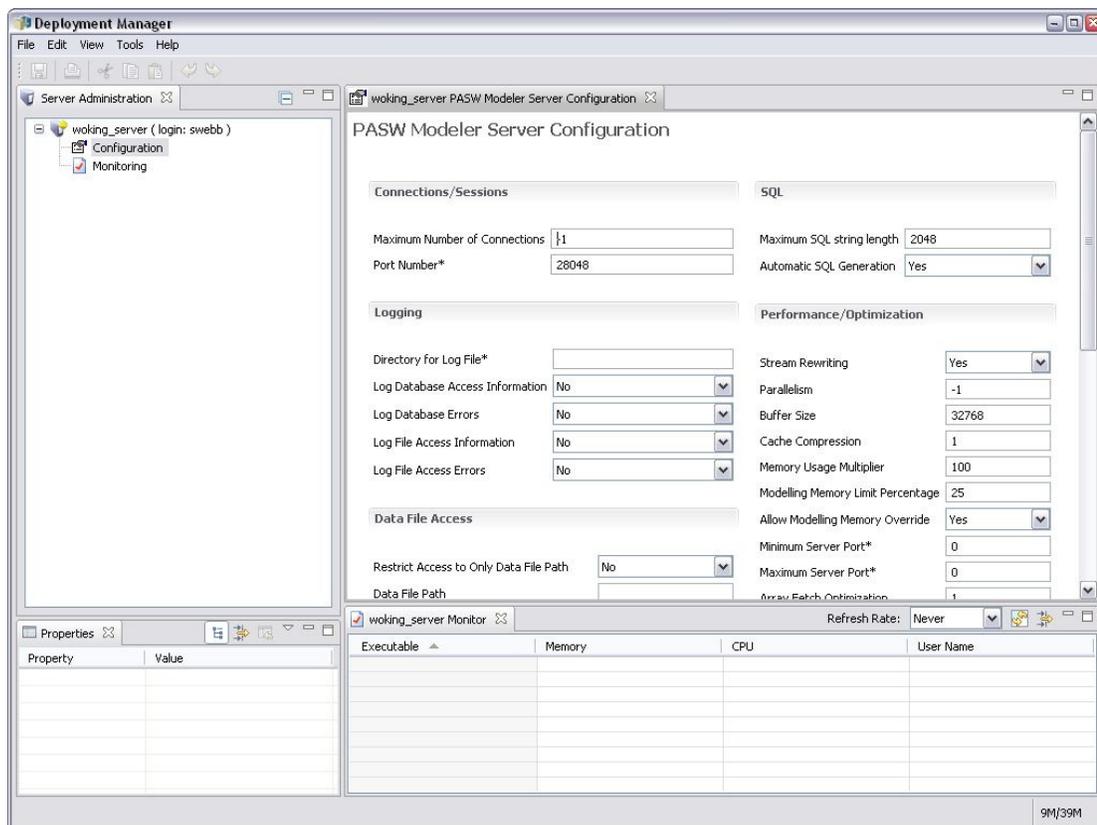
Muchas de las opciones disponibles a través de Modeler Administration Console también se pueden especificar en el archivo *options.cfg*, que se encuentra en el directorio de instalación de SPSS Modeler Server bajo */config*. Sin embargo, la aplicación Modeler Administration Console ofrece una interfaz gráfica compartida que permite conectarse, configurar y controlar varios servidores.

### **Inicie Modeler Administration Console**

- ▶ En el menú Inicio de Windows, seleccione [Todos los] Programas, a continuación, IBM SPSS Collaboration and Deployment Services, Deployment Manager y, por último, IBM SPSS Modeler Administration Console <nn.n>.

Cuando ejecute por primera vez la aplicación, verá los paneles Administración de servidores y Propiedades vacíos (a menos que ya haya instalado Deployment Manager con una conexión del servidor de IBM SPSS Collaboration and Deployment Services ya configurada). Después de configurar Modeler Administration Console, el panel del administrador del servidor que aparece a la izquierda mostrará un nodo para cada SPSS Modeler Server que desee administrar. El panel de la derecha muestra las opciones de configuración del servidor seleccionado. Primero debe [configurar una conexión](#) para cada servidor que desee administrar.

Figura 4-1  
Interfaz de Modeler Administration Console



## Configuración de acceso con Modeler Administration Console

El acceso del administrador a SPSS Modeler Server mediante Modeler Administration Console se controla con la línea `administrators` del archivo `options.cfg`, situado en el directorio de instalación de SPSS Modeler Server en `/config`. Esta línea está comentada por defecto, de manera que debe modificar esta línea para permitir el acceso a usuarios específicos, o utilizar `*` para permitir el acceso a todos los usuarios, tal y como se muestra en los ejemplos siguientes:

```
administrators, "*"
administrators, "jsmith,mjones,achavez"
```

- La línea debe comenzar por `administrators`, y las entradas deben escribirse entre comillas. Las entradas distinguen entre mayúsculas y minúsculas.
- Separe los identificadores de usuarios con comas.
- Para las cuentas de Windows, no utilice nombres de dominios.
- Utilice el asterisco con precaución. Permite que cualquiera con una cuenta de usuario válida para IBM® SPSS® Modeler Server (en la mayoría de los casos cualquiera de la red) pueda iniciar sesión y cambiar las opciones de configuración.

## **Conexiones de SPSS Modeler Server**

Debe especificar una conexión para cada SPSS Modeler Server de la red que quiera administrar. A continuación, debe iniciar sesión en cada servidor. A pesar de que la conexión del servidor se guarda entre sesiones de Modeler Administration Console, no ocurre así con las credenciales de inicio de sesión. Debe iniciar sesión cada vez que inicie Modeler Administration Console.

### **Para configurar una conexión de Server**

- ▶ Asegúrese de que se ha iniciado el servicio IBM® SPSS® Modeler Server.
- ▶ En el menú Archivo, elija Nuevo y, a continuación, Conexión de servidor administrado.
- ▶ En la primera página del asistente, escriba el nombre de la conexión del servidor. El nombre es para su propio uso y debería ser descriptivo, por ejemplo, *Servidor de producción*. Asegúrese de que se ha establecido Tipo a IBM SPSS Modeler Server administrado; a continuación pulse en Siguiente.
- ▶ En la segunda página, escriba el nombre de host o la dirección IP del servidor. Si el puerto es distinto al puerto por defecto, escriba el número de puerto. Pulse en Finalizar. La nueva conexión del servidor aparece en el panel del administrador del servidor.

Ahora, para realizar las tareas de administración, deberá iniciar sesión.

### **Para iniciar sesión en Server**

- ▶ En el panel del administrador del servidor, pulse dos veces en el servidor en el que desee iniciar sesión.
- ▶ En el cuadro de diálogo de inicio de sesión, escriba sus credenciales. (Utilice su cuenta de usuario para el host del servidor.) Pulse en Aceptar.

Si se produce un error de inicio de sesión con el mensaje No se pueden obtener los derechos de administrador en el servidor, puede que se deba a que el acceso del administrador no se haya configurado correctamente. [Si desea obtener más información, consulte el tema Configuración de acceso con Modeler Administration Console el p. 25.](#)

Si el inicio de sesión falla y aparece el mensaje No se ha podido conectar al servidor '<servidor>', asegúrese de que el ID y la contraseña de usuario son correctos y, a continuación, asegúrese de que se está ejecutando el servicio de SPSS Modeler Server. Por ejemplo, en Windows, vaya a Panel de control > Herramientas administrativas > Servicios y compruebe la entrada de IBM SPSS Modeler Server. Si la columna Estado no muestra Iniciado, seleccione esta línea en la pantalla y pulse en Iniciar; a continuación, vuelva a intentar el inicio de sesión.

Tras iniciar sesión en SPSS Modeler Server, aparecen bajo el nombre del servidor dos opciones, [Configuración](#) y [Supervisión](#). Pulse dos veces en una de estas opciones.

## Configuración de SPSS Modeler Server

El panel de configuración muestra las opciones de configuración de SPSS Modeler Server. Utilice este panel para cambiar las opciones como desee. Pulse en Guardar en la barra de herramientas para guardar los cambios. (Tenga en cuenta que si cambia el número de puerto deberá reiniciar el servidor para el cambio tenga efecto.)

Las opciones se describen en las secciones siguientes, con la línea correspondiente de *options.cfg* entre paréntesis para cada opción. Las opciones que están visibles sólo en *options.cfg* se describen al final de esta sección.

### Conexiones/Sesiones

Figura 4-2  
Configuración de Conexiones/Sesiones

Connections/Sessions	
Maximum Number of Connections	-1
Port Number*	28047

**Número máximo de conexiones.** (max\_sessions) Número máximo de usuarios conectados a la vez. El valor -1 indica que no hay límite.

**Número de puerto.** (port\_number) El número de puerto al que escucha SPSS Modeler Server. Cámbielo si otra aplicación ya utiliza el puerto por defecto (28047). Los usuarios finales deben conocer el número de puerto para poder utilizar SPSS Modeler Server.

### Registro

Figura 4-3  
Configuración de registro

Logging	
Directory for Log File*	
Log Database Access Information	No
Log Database Errors	No
Log File Access Information	No
Log File Access Errors	No

**Directorio del archivo de registro.** (log\_directory) Ubicación del archivo *messages.log*. Una cadena vacía significa [ruta de instalación del servidor]/log. Tenga en cuenta que debe utilizar barras diagonales. Si desea obtener más información sobre el contenido de este archivo de registro, consulte [Registro del servidor el p. 42](#).

**Registrar información de acceso a bases de datos.** (log\_database\_access) Añade información de acceso de la base de datos al archivo de registro.

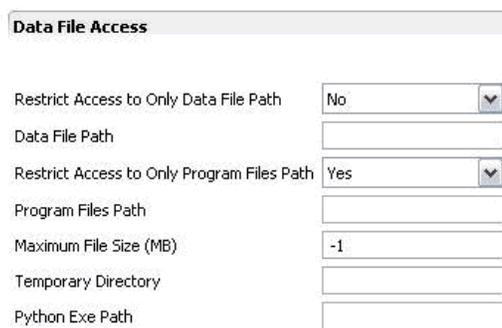
**Registrar errores de bases de datos.** (log\_database\_errors) Añade los errores de la base de datos al archivo de registro.

**Registrar información de acceso a archivos.** (log\_file\_access) Añade información de acceso a los archivos al archivo de registro.

**Registrar errores de acceso a archivos.** (log\_file\_errors) Añade los errores de acceso de los archivos al archivo de registro.

### Acceso a archivos de datos

Figura 4-4  
Configuración del acceso a archivos de datos



Data File Access	
Restrict Access to Only Data File Path	No
Data File Path	
Restrict Access to Only Program Files Path	Yes
Program Files Path	
Maximum File Size (MB)	-1
Temporary Directory	
Python Exe Path	

**Restringir el acceso sólo a la ruta de archivos de datos.** (data\_files\_restricted) Cuando se establece en *sí*, esta opción restringe los archivos de datos al directorio de datos estándar y a aquellos indicados en la ruta de archivos de datos que se indica a continuación.

**Ruta de archivos de datos.** (data\_file\_path) Una lista de directorios adicionales en los que se permite a los clientes leer y escribir archivos de datos. Esta opción se ignora a menos que la opción Restringir acceso sólo a la ruta del archivo de datos se active. Tenga en cuenta que debe utilizar barras diagonales en todos los nombres de ruta. Puede especificar varios directorios utilizando punto y coma (por ejemplo, [ruta de instalación del servidor]/data;c:data;c:temp).

**Restringir acceso sólo a la ruta de archivos de programas.** (program\_files\_restricted) Cuando se establece en *sí*, esta opción restringe el acceso a los archivos de programas al directorio estándar *bin* y a aquellos indicados en la ruta de archivos de programas a continuación.

**Ruta de archivos de programas.** (program\_file\_path) Lista de directorios adicionales desde los que se permite a los clientes ejecutar programas. Esta opción se ignora a menos que la opción Restringir acceso sólo a la ruta de archivos de programas se active. Tenga en cuenta que debe utilizar barras diagonales en todos los nombres de ruta. Puede especificar varios directorios utilizando punto y coma.

**Tamaño de archivo máximo (MB).** (max\_file\_size) Tamaño máximo de los archivos de datos temporales y exportados creados durante la ejecución de la ruta (no se aplica a los archivos de datos de SAS y SPSS Statistics). El valor -1 indica que no hay límite.

**Directorio temporal.** (temp\_directory) El directorio utilizado para almacenar archivos de datos temporales (archivos caché). Lo ideal sería que este directorio estuviese en un controlador o disco de alta velocidad diferente ya que la velocidad de acceso a este directorio puede tener

un impacto significativo en el rendimiento. Puede especificar varios directorios temporales, separándolos con comas. Estos se deberían ubicar en discos diferentes, la mayoría de las veces se utiliza el primer directorio, y se utilizan directorios adicionales para almacenar archivos de trabajo temporales cuando ciertas operaciones de preparación de datos (como la clasificación) utilizan paralelismo durante la ejecución. Cuando se permite que cada subproceso de ejecución utilice un disco diferente para el almacenamiento temporal se puede mejorar el rendimiento. Utilice barras diagonales en todas las especificaciones de ruta. *Nota:* Se generan archivos temporales en este directorio durante el inicio de SPSS Modeler Server. Asegúrese de que dispone de los derechos de acceso necesarios para este directorio (por ejemplo, si el directorio temporal es una carpeta de red compartida), de otro modo se producirá un error al iniciar SPSS Modeler Server.

**Ruta de acceso al ejecutable de Python.** (python\_exe\_path) Ruta de acceso completa al ejecutable de python incluido el nombre del ejecutable. Tenga en cuenta que es posible que Restringir acceso a sólo la ruta de archivos de programa deba establecerse en *No* dependiendo de la ubicación en la que se haya instalado python.

## Rendimiento/Optimización

Figura 4-5  
Configuración de rendimiento/optimización

Performance/Optimization	
Stream Rewriting	Yes
Parallelism	-1
Buffer Size	32768
Cache Compression	1
Memory Usage Multiplier	100
Modelling Memory Limit Percentage	25
Allow Modelling Memory Override	Yes
Minimum Server Port	0
Maximum Server Port	0
Array Fetch Optimization	1

**Reescritura de rutas.** (stream\_rewriting\_enabled) Permite al servidor optimizar rutas volviéndolas a escribir. Por ejemplo, el servidor debería distribuir operaciones de reducción de datos más cerca del nodo de origen para minimizar el tamaño del conjunto de datos lo más rápido posible. Normalmente se recomienda desactivar esta opción sólo si la optimización provoca un error o resultados inesperados. Esta configuración reemplaza la configuración de optimización de cliente correspondiente.

**Paralelismo.** (max\_parallelism) Describe el número de subprocesos de trabajo en paralelo que se permite utilizar a SPSS Modeler al ejecutar una ruta. Si se establece en 0 o en cualquier número negativo IBM® SPSS® Modeler hace coincidir el número de subprocesos con el número de procesadores disponibles en el ordenador, el valor por defecto para esta opción es -1. Para desactivar el proceso paralelo (para equipos con varios procesadores), establezca esta opción en 1. Para permitir un procesamiento paralelo limitado, establézcalo en un número inferior al

número de procesadores de su equipo. Tenga en cuenta que un procesador con doble núcleo o con subprocesos múltiples se trata como dos procesadores.

**Tamaño del búfer (bytes).** (`io_buffer_size`) Los archivos de datos transferidos desde el servidor al cliente se pasan por un búfer de este número de bytes.

**Compresión de caché.** (`cache_compression`) Un valor entero entre 0 y 9 que controla la compresión de caché y otros archivos en el directorio temporal del servidor. La compresión reduce la cantidad utilizada de espacio en disco, que puede ser importante cuando el espacio es limitado, y mejora el rendimiento reduciendo la cantidad de actividad de disco necesaria para crear y leer la caché. Además aumenta el tiempo de procesador pero casi siempre mediante la reducción del tiempo de acceso a disco. Observe que únicamente se pueden comprimir ciertas memorias caché, aquellas a las que se accede de manera secuencial. Esta opción no se aplica las cachés de acceso aleatorio, como las que utilizan los algoritmos de entrenamiento de redes. Un valor de 0 desactiva la compresión por completo. Los valores de 1 en adelante proporcionan grados crecientes de compresión, pero con un coste correspondiente en tiempo de acceso. El valor por defecto es 1; los valores superiores no suelen ser necesarios excepto cuando el espacio en disco es muy importante.

**Multiplicador del uso de memoria.** (`memory_usage`) Controla la proporción de memoria física asignada para ordenar y realizar otras funciones en cachés de memoria. El valor por defecto es 100, que se corresponde con un 10% de memoria física aproximadamente. Aumente este valor para mejorar el rendimiento de clasificación donde la memoria libre está disponible, pero hágalo con precaución: si lo aumenta mucho puede provocar una paginación excesiva.

**Porcentaje del límite de memoria de modelado.** (`modelling_memory_limit_percentage`) Controla la proporción de memoria física asignada para los modelos de  $k$ -medias, Kohonen y red neuronal de entrenamiento. El valor por defecto es 25%. Aumente este valor para mejorar el rendimiento de entrenamiento donde la memoria libre está disponible, pero hágalo con precaución: si lo aumenta mucho puede provocar una paginación excesiva al volcar datos en el disco.

**Permitir reemplazo de memoria de modelado.** (`allow_modelling_memory_override`) Activa o desactiva la opción Optimizar velocidad en ciertos nodos de modelado. El valor por defecto está activado. Esta opción permite al algoritmo de modelado reclamar toda la memoria disponible, omitiendo la opción de límite de porcentaje. Puede que desee desactivar esto si necesita compartir recursos de memoria en el equipo servidor.

**Puerto del servidor máximo y mínimo.** (`max_server_port` y `min_server_port`) Especifica el rango de números de puerto que se pueden utilizar para las conexiones de los sockets adicionales entre cliente y servidor necesarios para modelos interactivos y ejecución de rutas. Esto requiere que el servidor esté en línea en otro puerto; si no se restringe el rango, los usuarios en sistemas con cortafuegos podrían tener problemas. El valor por defecto para ambos es  $-1$ , lo que indica “no restricción.” Así, por ejemplo, para establecer el servidor en línea en un puerto 8000 o superior, definiría `min_server_port` a 8000 y `max_server_port` a  $-1$ .

Tenga en cuenta que debe abrir puertos adicionales en el puerto del servidor principal para abrir o ejecutar una ruta y más puertos de forma correspondiente si desea abrir o ejecutar rutas concurrentes. Esto es necesario para capturar información a partir de la ejecución de la ruta.

Por defecto, SPSS Modeler utilizará cualquier puerto abierto que esté disponible; si no encuentra ninguno (por ejemplo, si están todos cerrados por un cortafuegos), aparecerá un error cuando ejecute la ruta. Para configurar el intervalo de puertos, SPSS Modeler necesitará dos puertos abiertos (junto con el puerto principal del servidor) disponibles por ruta concurrente, junto

con un puerto adicional para cada cliente conectado que tenga una conexión ODBC. Para ejecutar más rutas simultáneas, puede ampliar el rango de números de puerto y abrir más puertos en el cortafuegos correctamente.

*Nota:* Si cambia estos parámetros, deberá reiniciar SPSS Modeler Server para que el cambio surta efecto.

**Optimización de obtención de matrices.** (`sql_row_array_size`) Controla la forma en que SPSS Modeler Server obtiene datos del origen de datos ODBC. El valor por defecto es 1, que obtiene una sola fila cada vez. Al aumentar el valor el servidor lee la información en fragmentos mayores, obteniendo el número especificado de filas en una matriz. Con algunas combinaciones de sistemas operativos/bases de datos, esto podría producir mejoras en el rendimiento de instrucciones SELECT.

## SQL

Figura 4-6  
Configuración de SQL

SQL

Maximum SQL string length

Automatic SQL Generation

**Longitud de cadena de SQL máxima.** (`max_sql_string_length`) Longitud máxima de una cadena importada de una base de datos con SQL. Los valores de cadena con más caracteres se truncan por la derecha sin advertencias. El rango válido es de 1 a 65.535 caracteres.

**Generación de SQL automática.** (`sql_generation_enabled`) Permite la generación automática de SQL para rutas, lo que puede mejorar significativamente el rendimiento. El valor por defecto está activado. Se recomienda desactivar esta opción sólo si la base de datos no puede admitir las consultas enviadas desde SPSS Modeler Server. Tenga en cuenta que esta configuración sobrescribe la configuración de optimización de clientes correspondiente; tenga en cuenta que también con fines de puntuación, la generación de SQL se debe activar por separado para cada nodo de modelado, independientemente de esta configuración.

## SSL

Figura 4-7  
Configuración de SSL

SSL

Enable SSL

Certificate File

Private Key File

Private Key Password

**Activar SSL.** (`ssl_enabled`) Activa el cifrado SSL para las conexiones establecidas entre SPSS Modeler y SPSS Modeler Server.

**Archivo de certificado.** (ssl\_certificate\_file) El archivo de certificado SSL público que se cargará cuando se inicie el servidor.

**Archivo de clave privada.** (ssl\_private\_key\_file) El archivo de clave SSL privada que se cargará cuando se inicie el servidor.

**Contraseña de clave privada.** (ssl\_private\_key\_password) Clave de contraseña para el certificado especificado. Tenga en cuenta que esta clave debe estar firmada al menos por una de las entidades emisoras de certificados que aparecen en el archivo de certificado configurado.

### Configuración de Coordinator of Processes

Figura 4-8  
Valores de configuración de Coordinator of Processes

Co-ordinator Of Processes Configuration	
Host*	spsscop
Port Number*	8080
Login Name*	
Password*	
Enabled*	No
Server Name*	
Description*	
Update Interval (min)*	2
Weight	1
Service Host*	

Items marked (\*) require a server restart to take effect.

**Host.** (cop\_host) El nombre de host o dirección IP del servicio de Coordinator of Processes. El nombre “spsscop” por defecto es un nombre redundante que los administradores pueden añadir como alias para el host de IBM SPSS Collaboration and Deployment Services en DNS.

**Número de puerto.** (cop\_port\_number) El número de puerto del servicio de Coordinator of Processes. El valor por defecto, 8080, es el valor por defecto de IBM SPSS Collaboration and Deployment Services.

**Nombre de inicio de sesión.** (cop\_user\_name) El Nombre de usuario para la autenticación del servicio de Coordinator of Processes. Es un nombre de inicio de sesión de IBM SPSS Collaboration and Deployment Services así que puede incluir un prefijo de proveedor de seguridad (por ejemplo: ad/jsmith).

**Contraseña.** (cop\_password) La contraseña cifrada para la autenticación del servicio de Coordinator of Processes.

Para crear el archivo de contraseña, necesitará utilizar el programa de utilidad de contraseña, *pwutil*, que se encuentra en el directorio *bin* de la instalación de SPSS Modeler Server.

**Activado.** (cop\_enabled) Determina si el servidor debe intentar registrarse con Coordinator of Processes. El valor por defecto *no* debe registrarse porque el administrador debe seleccionar qué servicios deben publicitarse a través de Coordinator of Processes.

**Nombre de servidor.** (cop\_service\_name) El nombre de esta instancia de SPSS Modeler Server; el valor por defecto es el nombre de host.

**Descripción.** (cop\_service\_description) Una descripción de esta instancia.

**Update interval (min) [Intervalo de actualización (min)].** (cop\_update\_interval) El número de minutos entre mensajes activados; el valor por defecto es 2.

**Ponderación.** (cop\_service\_weight) La ponderación de esta instancia, especificada como un número entero entre 1 y 10. Una mayor ponderación atrae más conexiones. El valor por defecto es 1.

**Service host (Host de servicio).** (cop\_service\_host) El nombre de host totalmente aceptado del host de IBM® SPSS® Modeler Server. El valor por defecto del nombre de host se deriva automáticamente; el administrador puede sustituir el valor por defecto de hosts con interfaz múltiple.

### ***Opciones visibles únicamente en options.cfg***

**default\_sql\_string\_length.** Especifica el ancho por defecto de columnas de cadena que se crearán dentro de las tablas de caché de la base de datos. Si no hay información de tipo anterior, se crearán campos de cadenas en las tablas de caché de la base de datos con un ancho por defecto de 255. Si tiene valores más anchos que esto en sus datos, instancie un nodo Tipo anterior con esos valores o establezca este parámetro con un valor que sea lo suficientemente grande para admitir esos valores de cadena.

**max\_transfer\_size.** Sólo para uso interno del sistema; no modificar.

**start\_process\_as\_login\_user.** Establezca esto como true si utiliza un inicio de sesión único y está conectado a un servidor de IBM® SPSS® Collaboration and Deployment Services que no utiliza el mismo método de autenticación que el host de IBM® SPSS® Modeler Server. La configuración de esta opción como true permite que los procesos secundarios utilicen los detalles de inicio de sesión de SPSS Modeler Server.

**administrators.** Especifica los nombres de usuario de aquellos usuarios a los que desea conceder acceso de administrador. [Si desea obtener más información, consulte el tema Configuración de acceso con Modeler Administration Console el p. 25.](#)

**supported\_sql\_type\_logging.** Sólo para fines de depuración; no modificar a menos que un representante del servicio de asistencia técnica indique lo contrario.

**allow\_config\_custom\_overrides.** No modificar a menos que un representante del servicio de asistencia técnica indique lo contrario.

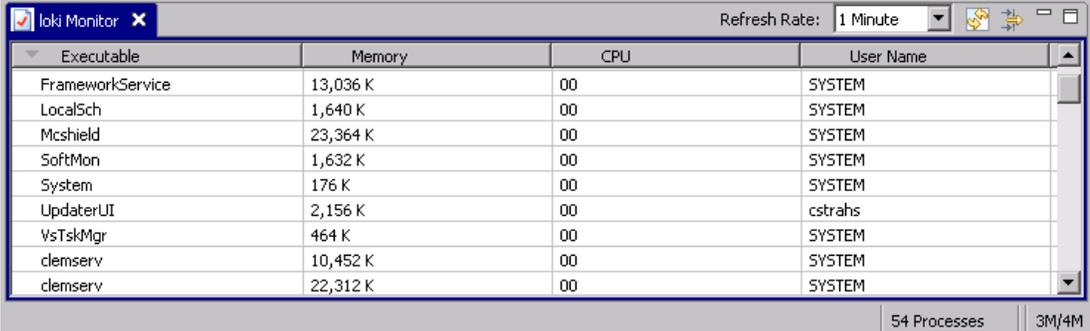
## ***Supervisión de SPSS Modeler Server***

El panel de supervisión de Modeler Administration Console muestra una instantánea de todos los procesos que se ejecutan en el ordenador de SPSS Modeler Server, parecida a la del Administrador de tareas de Windows. Para activar el panel de supervisión, pulse dos veces en el nodo de supervisión bajo el servidor deseado en el panel del administrador del servidor. Aparecerá en el

panel una instantánea actual con los datos del servidor. Los datos se actualizan a la velocidad indicada (un minuto por defecto). Para actualizar los datos manualmente, pulse en el botón Actualizar. Para mostrar en esta lista sólo los procesos de SPSS Modeler Server, pulse en el botón Filtrar los procesos ajenos a SPSS Modeler.

Figura 4-9

Panel de supervisión de IBM SPSS Modeler Administration Console



Executable	Memory	CPU	User Name
FrameworkService	13,036 K	00	SYSTEM
LocalSch	1,640 K	00	SYSTEM
Mcshield	23,364 K	00	SYSTEM
SoftMon	1,632 K	00	SYSTEM
System	176 K	00	SYSTEM
UpdaterUI	2,156 K	00	cstrahs
VsTskMgr	464 K	00	SYSTEM
clmserv	10,452 K	00	SYSTEM
clmserv	22,312 K	00	SYSTEM

Refresh Rate: 1 Minute

54 Processes 3M/4M

## Uso del archivo *options.cfg*

El archivo de texto *options.cfg* se encuentra en el directorio *[ruta de instalación del servidor]/config*. Cada parámetro de configuración está representado por un par nombre-valor separado por coma, donde el **nombre** es el nombre de la opción y el **valor**, su valor. El signo almohadilla (#) señala un comentario.

*Nota:* A excepción de lo siguiente, todas las opciones de configuración se pueden cambiar mediante IBM® SPSS® Collaboration and Deployment Services Deployment Manager, en vez de con este archivo de configuración:

- administrators
- allow\_custom\_config\_overrides
- max\_transfer\_size
- start\_process\_as\_login\_user
- supported\_sql\_type\_logging

Con Deployment Manager, podrá evitar que el servidor se reinicie para todas las opciones menos el puerto del servidor. [Si desea obtener más información, consulte el tema Uso de IBM SPSS Modeler Administration Console el p. 24.](#)

### **Opciones de configuración que se pueden añadir al archivo por defecto**

Por defecto, el almacenamiento en caché en la base de datos está habilitado para IBM® SPSS® Modeler Server. Puede deshabilitar esta función añadiendo la siguiente línea al archivo *options.cfg*:

```
enable_database_caching, N
```

Al realizar esta acción, los archivos temporales se crearán en el servidor y no en la base de datos.

**Para ver o cambiar las opciones de configuración de IBM SPSS Modeler Server**

- ▶ Abra el archivo *options.cfg* con un editor de texto.
- ▶ Busque las opciones que le interesen. Para obtener una lista completa de opciones, consulte [Configuración de SPSS Modeler Server el p. 27](#).
- ▶ Edite los valores como considere oportuno. Tenga en cuenta que todos los valores de nombre de ruta deben usar como separador una barra diagonal (/) y no un carácter de barra invertida.
- ▶ Guarde el archivo.
- ▶ Detenga SPSS Modeler Server y vuelva a iniciarlo de manera que los cambios tengan efecto. [Si desea obtener más información, consulte el tema Inicio y parada de IBM SPSS Modeler Server el p. 21](#).

**Cierre de conexiones de base de datos sin utilizar**

Por defecto, IBM® SPSS® Modeler almacena en la caché al menos una conexión a una base de datos una vez que se ha accedido a dicha conexión. La sesión de la base de datos se mantiene abierta aunque no se ejecuten las rutas que necesitan acceso a la base de datos.

El almacenamiento en caché de las conexiones de la base de datos pueden mejorar los tiempos de ejecución al eliminar la necesidad de que SPSS Modeler se vuelva a conectar a la base de datos cada vez que se ejecuta una ruta. Sin embargo, en algunos entornos, es importante que las aplicaciones liberen recursos de la base de datos con la mayor rapidez posible. Si demasiadas sesiones de SPSS Modeler mantienen conexiones a la base de datos que ya no se utilizan, se pueden agotar los recursos de la base de datos.

Para evitar esta posibilidad, desactive la opción SPSS Modeler de *cache\_connection* en un archivo de configuración de base de datos personalizada. Al llevarlo a cabo, también puede hacer que SPSS Modeler sea más resistente a los errores en la conexión de base de datos (como los tiempos de espera) que se pueden producir cuando una sesión de SPSS Modeler usa las conexiones durante un período de tiempo prolongado.

Para provocar el cierre de las conexiones de base de datos no utilizadas:

- ▶ Localice el directorio *[ruta de instalación del servidor]/config*.
- ▶ Añada el archivo siguiente (o ábralo si ya existe):  
*odbc-custom-properties.cfg*
- ▶ Añada la siguiente línea al archivo:  
`cache_connection, N`
- ▶ Guarde y cierre el archivo.
- ▶ Reinicie IBM® SPSS® Modeler Server para que los cambios surtan efecto.

## **Ubicación de licencia de IBM SPSS Statistics**

Para permitir que IBM® SPSS® Modeler utilice los nodos Transformación de Statistics, Modelo de Statistics y Resultados de Statistics, debe tener instalada una copia de IBM® SPSS® Statistics con licencia en el ordenador donde se ejecute la ruta. Cuando realice la ejecución con respecto a un servidor, la configuración de licencia se deberá realizar en el servidor.

### **Windows**

- ▶ A partir del símbolo de sistema, cambie el directorio bin de IBM® SPSS® Modeler Server y ejecute:

```
statisticsutility -location =<ruta al archivo de licencia de IBM SPSS Statistics Server>
```

Donde <path to SPSS Statistics Server license file> es el directorio de instalación de un servidor SPSS Statistics con licencia.

### **UNIX**

- ▶ A partir del símbolo de sistema, cambie el directorio bin de SPSS Modeler Server y ejecute:

```
./statisticsutility -location =<ruta al archivo de licencia de IBM SPSS Statistics Server>
```

Donde <path to SPSS Statistics Server license file> es el directorio de instalación de un servidor SPSS Statistics con licencia.

## **Uso de SSL para transferencias de datos seguras**

SSL (Security Sockets Layer) es un protocolo para cifrar datos transferidos entre dos ordenadores. SSL garantiza que la comunicación entre los ordenadores es segura. SSL puede cifrar la autenticación de un nombre de usuario/contraseña y el contenido de un intercambio entre un servidor y un cliente.

### **Cómo funciona SSL**

SSL se basa en las claves públicas y privadas del servidor, junto con un certificado de clave pública que enlaza la identidad del servidor con su clave pública.

- ▶ Cuando un cliente se conecta a un servidor, el cliente autentifica el servidor con el certificado de clave pública.
- ▶ El cliente genera a continuación un número aleatorio, cifra el número con la clave pública del servidor, y envía el mensaje cifrado de vuelta al servidor.
- ▶ El servidor descifra el número aleatorio con su clave privada.
- ▶ A partir del número aleatorio, tanto el servidor como el cliente crean las claves de sesión utilizadas para cifrar y descifrar la información posterior.

El certificado de clave pública está firmado normalmente por una entidad emisora de certificados. Las entidades emisoras de certificados, como VeriSign y Thawte, son organizaciones que emiten, autentifican y gestionan las credenciales de seguridad contenidas en los certificados de clave pública. Esencialmente, la entidad emisora de certificados confirma la identidad del servidor. La entidad emisora de certificados normalmente cobra una tarifa por un certificado, pero también pueden generarse certificados autofirmados.

## **Comunicaciones cliente-servidor y servidor-servidor seguras con SSL**

Los pasos principales en las comunicaciones cliente-servidor y servidor-servidor seguras con SSL son:

- ▶ Instalar OpenSSL en el ordenador servidor.
- ▶ Obtener e instalar el certificado SSL y las claves.
- ▶ Activar y configurar SSL en la aplicación de administración de servidores (IBM® SPSS® Collaboration and Deployment Services Deployment Manager).
- ▶ Si se desea, instalar cifrado de potencia ilimitada en los ordenadores cliente.
- ▶ Si se utiliza un certificado autofirmado, copiar el certificado en el ordenador cliente.
- ▶ Añadir el certificado al almacén de claves del cliente.
- ▶ Indicar a los usuarios finales que activen SSL cuando se conecten con el servidor.

*Nota:* En algunas ocasiones un servidor actúa como cliente. Un ejemplo es IBM® SPSS® Statistics Server conectándose al IBM® SPSS® Collaboration and Deployment Services Repository. En este caso, SPSS Statistics Server es el *cliente*.

### **Instalar OpenSSL**

Si OpenSSL no está aún disponible en el servidor, deberá instalarlo.

- ▶ Descargue OpenSSL desde <http://www.openssl.org/>. Asegúrese de utilizar la versión de OpenSSL apropiada para la versión del servidor:

Tabla 4-1  
Versiones de OpenSSL

Producto servidor	Versión de OpenSSL compatible
IBM® SPSS® Statistics 17-19 (no Linux® on System z®) IBM® SPSS® Modeler 14.2 (no Linux® en System z®)	0.9.8 y sus subversiones (0.9.8a, 0.9.8b, etc.)
SPSS Statistics 19 (Linux® on System z®) SPSS Modeler 14.2 (Linux® en System z®)	1.0.0

Nota: IBM® SPSS® Modeler Server para UNIX, cuando se configura para SSL, busca vincular dos bibliotecas con nombres como (.so puede ser específico de la plataforma):

- *libssl.so.0.9.8*
- *libcrypto.so.0.9.8*

Si utiliza una subversión de SSL (0.9.8a, 0.9.8b, etcétera), debe crear vínculos simbólicos para estas bibliotecas con las versiones *xxx.0.9.8* que espera SPSS Modeler Server.

- ▶ Siga las instrucciones para instalar y configurar el software. Si construye OpenSSL por sí mismo, tenga en cuenta las siguientes directrices:

**Windows.** OpenSSL debe construirse con bibliotecas DLL (que son multiproceso por defecto).

**UNIX.** OpenSSL debe admitir subprocessos múltiples (que no es siempre el valor por defecto) y bibliotecas compartidas.

- ▶ Asegúrese de que el sistema puede localizar las bibliotecas SSL:

**Windows.** Asegúrese de que los módulos OpenSSL están incluidos en la ruta del sistema.

**UNIX.** Agregue las bibliotecas SSL al final de la ruta de la biblioteca de carga definida en *modelersrv.sh*. Por ejemplo, en Solaris, añada esta línea al proceso de inicio del servidor:

```
LD_LIBRARY_PATH_64=$LD_LIBRARY_PATH_64:/usr/local/openssl; export LD_LI  
BRARY_PATH_64
```

### ***Obtención e instalación de certificados y claves SSL***

- ▶ Obtenga un archivo de certificado y de clave SSL. Existen dos formas de hacerlo:
  - Adquirirlos a través de una entidad emisora de certificados pública (como Verisign o Thawte). La entidad emisora de certificados pública firma el certificado para verificar el servidor que lo utiliza.
  - Genera los archivos de clave y de certificado mediante una entidad emisora de certificados autofirmados interna. OpenSSL proporciona una herramienta de administración de certificados para este fin, o puede buscar instrucciones en Internet acerca de la creación de un certificado autofirmado de SSL.
- ▶ Copie el archivo de certificado y de clave a un directorio local en el servidor. Las claves públicas y privadas pueden almacenarse en un archivo. Asegúrese de que los permisos de acceso deniegan la apertura de este directorio.
- ▶ Instale el certificado SSL y las claves en el servidor de aplicaciones. Para obtener información adicional sobre cómo las claves y los certificados interactúan con un servidor de aplicaciones específico, consulte la documentación original del proveedor. Tenga en cuenta que es posible que se le solicite añadir el certificado y las claves al almacén de claves Java.

### ***Active y configure SSL en el IBM SPSS Collaboration and Deployment Services Deployment Manager***

- ▶ Inicie la aplicación de administración del servidor (IBM® SPSS® Collaboration and Deployment Services Deployment Manager) y conecte con el servidor.

- ▶ En la página de configuración, asigne Secure Sockets Layer a Yes.
- ▶ En el archivo de clave pública SSL, especifique la ruta completa al archivo de clave pública.
- ▶ En el archivo de clave privada SSL, especifique la ruta completa al archivo de clave privada.

*Nota:* Si las claves pública y privada están almacenadas en un archivo, especifique el mismo archivo en el archivo de clave pública SSL y en el archivo de clave privada SSL.

- ▶ Elija en los menús:  
File > Guardar
- ▶ Reinicie el servicio o centinela servidor. Cuando reinicie, se le solicitará la contraseña SSL. En Windows, puede seleccionar Recordar contraseña para almacenar la contraseña de forma segura. Esta opción elimina la necesidad de introducir la contraseña cada vez que se inicia el servidor.

### **Configuración de SSL en IBM SPSS Modeler Server (sólo sistemas UNIX)**

Si se necesita una contraseña para acceder al archivo de clave SSL privada, en los sistemas UNIX puede proporcionar la contraseña de una de las maneras siguientes:

- Especifíquela en el archivo *options.cfg* mediante la opción `ssl_private_key_password`. [Si desea obtener más información, consulte el tema SSL el p. 31.](#)
- Añada un modificador al archivo *modelersrv.sh* para pedir la contraseña al inicio.

Si utiliza esta segunda opción, busque la siguiente línea en el archivo *modelersrv.sh*:

```
if "$INSTALLEDPATH/$SCLEMDNAME" -server $ARGS; then
```

Añada el modificador `-request_ssl_password` de la manera siguiente:

```
if "$INSTALLEDPATH/$SCLEMDNAME" -request_ssl_password -server $ARGS; then
```

### **Instalación de cifrado de potencia ilimitada**

Java Runtime Environment, que se incluye como parte del producto, tiene activado el cifrado de potencia para exportación de EE.UU. Para mejorar la seguridad de los datos, recomendamos que se actualice a cifrado de potencia ilimitada.

- ▶ Descargue Java Cryptography Extension (JCE) Unlimited Strength Jurisdiction Policy Files 5.0 de <http://java.sun.com/javase/downloads/index.jsp>.
- ▶ Descomprima el archivo descargado.
- ▶ Copie los dos archivos *.jar*, *local\_policy.jar* y *US\_export\_policy.jar*, en *<carpeta de instalación>/jre/lib/security*, donde *<carpeta de instalación>* es la carpeta en la que ha instalado el producto.

### **Adición del certificado al almacén de claves del cliente (para conexiones al repositorio de o IBM SPSS Modeler Server)**

*Nota:* Omite este paso si está utilizando un certificado que está firmado por una entidad emisora de certificados.

Si está utilizando SSL para conectar con el repositorio de IBM® SPSS® Collaboration and Deployment Services o IBM® SPSS® Modeler Server y está utilizando certificados autofirmados, tendrá que añadir el certificado al almacén de claves Java del cliente. Los siguientes pasos se realizan en el ordenador *cliente*.

Tenga en cuenta que un ordenador servidor también puede actuar como un cliente. Un ejemplo es IBM® SPSS® Statistics Server conectándose al IBM® SPSS® Collaboration and Deployment Services Repository. En este caso, SPSS Statistics Server es el *cliente*, y por lo tanto tiene que copiar el certificado para el servidor del IBM SPSS Collaboration and Deployment Services Repository a SPSS Statistics Server.

- ▶ Copie el archivo de certificado (.cer) del servidor de SSL al directorio `<install_dir>\jre\lib\security` en el equipo cliente.
- ▶ En el mismo directorio, haga clic con el botón derecho del ratón en el archivo *cacerts* y seleccione Properties.
- ▶ Cancele la selección de la casilla de verificación Sólo lectura.
- ▶ En el símbolo del sistema, cambie los directorios a la siguiente ubicación, donde `<product install dir>` es el directorio en el que ha instalado el producto:

```
<product install dir>/jre/bin
```

- ▶ Escriba el siguiente comando:

```
keytool -import -alias <nombre de alias> -file <ruta al certificado> -keystore <ruta al almacén de cla
```

Donde `<nombre de alias>` es un alias arbitrario del certificado, `<ruta al certificado>` es la ruta del certificado y `<ruta al almacén de claves>` es la ruta completa del almacén de claves de Java, que puede ser `<product install dir>/lib/security/jssecacerts` o `<product install dir>/lib/security/cacerts`.

- ▶ Cuando se le solicite, introduzca la contraseña del almacén de claves, que es por defecto, `changeit`.
- ▶ Cuando se le pregunte si confía en el certificado, introduzca `yes`.

### **Indicaciones de la activación de SSL a usuarios finales**

Cuando los usuarios finales conecten con el servidor a través de un producto cliente, tendrán que activar SSL en el cuadro de diálogo para conectar con el servidor. Asegúrese de indicar a sus usuarios que seleccionen la casilla de verificación apropiada. Por ejemplo en IBM® SPSS® Modeler, los usuarios deben seleccionar la casilla de verificación Asegurar conexión segura en el Repositorio: Cuadro de diálogo del servidor, accesible desde la opción Repositorio en el menú Herramientas.

## Configuración de prefijo URL

Si IBM® SPSS® Collaboration and Deployment Services Repository está configurado para acceso SSL, el valor de Prefijo URL debe modificarse de la manera siguiente:

1. Inicie sesión en el repositorio utilizando la consola basada en navegador.
2. Abra la opción de configuración *Prefijo URL*.  
Configuración > Configuración > Prefijo URL
3. Establezca el valor del prefijo a `https` en vez de `http` y establezca el valor de puerto al número de puerto SSL. Por ejemplo:

[Por defecto]

`http://<nombrehost>:<puerto>`

[Con SSL]

`https://<nombrehost>:<puertoSSL>`

## Seguridad LDAP con SSL

LDAP (Lightweight Directory Access Protocol) es un estándar IETF (Internet Engineering Task Force) para intercambiar información entre directorios y bases de datos de red que contengan cualquier nivel de información. Los sistemas que requieren seguridad adicional, proveedores LDAP, como Microsoft Active Directory, pueden funcionar sobre SSL (Secure Sockets Layer), siempre que el servidor o la aplicación Web admitan LDAP sobre SSL. El uso de SSL en conjunción con LDAP puede garantizar que las contraseñas de inicio de sesión, la información de aplicaciones y otros datos confidenciales no se ven comprometidos o puedan robarse.

El siguiente ejemplo ilustra cómo activar LDAPS utilizando Microsoft Active Directory como proveedor de seguridad. Para obtener información más específica sobre cualquiera de estos pasos o para localizar detalles para una versión en particular del proveedor de seguridad, consulte la documentación original del proveedor.

1. Verifique que Active Directory y Enterprise Certificate Authority están instalados y funcionando.
2. Utilice la entidad emisora de certificados para generar un certificado, e importe el certificado en el almacén de certificados de la instalación de IBM® SPSS® Collaboration and Deployment Services Deployment Manager. Esta acción permitirá que se establezca la conexión LDAPS entre IBM® SPSS® Collaboration and Deployment Services Repository y un servidor Active Directory.

Para configurar Deployment Manager para realizar conexiones Active Directory seguras, verifique que existe una conexión en el repositorio.

3. Inicie la IBM® SPSS® Collaboration and Deployment Services Deployment Manager.
4. En el menú Herramientas, seleccione Administración de servidor.
5. Inicie sesión en un servidor administrado previamente definido.
6. Pulse dos veces en el icono Configuración para que el servidor amplíe la jerarquía.
7. Pulse dos veces en el icono Proveedores de seguridad para que se amplíe la jerarquía.

8. Pulse dos veces en el proveedor de seguridad Active Directory.
9. Introduzca los valores de configuración para la instancia de Active Directory con los certificados de seguridad instalados.
10. Seleccione la casilla de verificación Utilizar SSL.
11. Tenga en cuenta el nombre en el campo Usuario de dominio. Los inicios de sesión posteriores utilizando Active Directory están autenticados utilizando SSL.

Para obtener información adicional sobre la instalación, configuración e implementación de LDAPS en un servidor de aplicaciones en particular, consulte la documentación original del proveedor.

## **Registro del servidor**

IBM® SPSS® Modeler Server mantiene un registro de las acciones importantes en un archivo de registro denominado *messages.log*. Por defecto, este archivo se encuentra en *[ruta de instalación del servidor]/log*, si bien esta ubicación es configurable. Cuando SPSS Modeler Server se inicia, el archivo de registro se actualiza para reflejar la hora de inicio y el número de puerto al que escucha. A partir de ahí, el archivo de registro se actualiza con el resultado siempre que un cliente intente conectarse al servidor.

El servidor registra los siguientes eventos por defecto en el archivo de registro:

- Servidor iniciado
- Servidor detenido
- Conexión aceptada
- Conexión rechazada
- Sesión iniciada
- Sesión detenida
- Error del sistema
- Inicio de sesión correcto
- Error de inicio de sesión

Si desea registrar información sobre los eventos de acceso a la base de datos, cambie Registrar información de acceso a bases de datos y Registrar errores de bases de datos a *sí* en la configuración de SPSS Modeler Server. De este modo, se registrarán los siguientes eventos de acceso a base de datos:

- Inicio de sesión correcto en base de datos
- Error de inicio de sesión en base de datos
- Cierre de sesión de base de datos

Si desea registrar información sobre los eventos de acceso al archivo, cambie Registrar información de acceso a archivos y Registrar errores de acceso a archivos a *sí* en la configuración de SPSS Modeler Server. De este modo, se registrarán los siguientes eventos de acceso a archivo:

- Abrir/crear archivo
- Cerrar archivo

A continuación se muestra un ejemplo de archivo de registro (se han añadido números de línea):

```
1. 2006/06/14 08:41:23 [167]: IBM SPSS Modeler Server Version 0.1.0 started
2. 2006/06/14 08:41:23 [167]: Listening on port 28028
3. 2006/06/14 08:49:28 [167]: Accepted connection from localhost:1827
4. 2006/06/14 08:49:28 [167]: Session 105 (localhost:1827) started
5. 2006/06/14 08:49:29 [167-105]: Login failed for user: SPSS\awilson
6. 2006/06/14 09:27:27 [167]: Session 105 ended
7. 2006/06/14 09:53:39 [167]: Accepted connection from localhost:1837
8. 2006/06/14 09:53:39 [167]: Session 63 (localhost:1837) started
9. 2006/06/14 09:53:40 [167-63]: Login succeeded for user: SPSS\awilson
10. 2006/06/14 10:27:31 [167]: Session 63 ended
```

Todos los elementos del archivo de registro contienen una marca de tiempo. Las líneas 1 y 2 se añadieron al archivo de registro cuando se inició SPSS Modeler Server. La línea 3 se añadió cuando un usuario del equipo *hostlocal* en el puerto número *1827* intentó conectarse al servidor. La línea 4 indica que se inició una sesión de servidor con este usuario. La línea 5 refleja que el usuario no pudo iniciar sesión correctamente. Más tarde, el usuario finalizó la sesión. Media hora más tarde, se intentó establecer otra conexión desde *hostlocal* y se inició una sesión. Esta vez el inicio de sesión se realizó con éxito. Transcurrida otra media hora, la sesión finalizó y se cerró la conexión.

## ***Conceptos básicos de rendimiento***

El rendimiento real cuando se analizan datos se ve afectado por una serie de factores, que abarcan desde la configuración de la base de datos y el servidor hasta el ordenamiento de nodos individuales dentro de una ruta. En general, puede obtener el mejor rendimiento realizando lo siguiente:

- Almacenar los datos en un DBMS y utilizar la optimización de SQL cuando sea posible.
- Utilizar hardware que cumpla o supere las recomendaciones proporcionadas en [el capítulo 2](#).
- Asegurarse de que la configuración de rendimiento y optimización del servidor se ha establecido correctamente. Algunos de estos valores de configuración también están disponibles en el cliente, pero sólo son aplicables al utilizar IBM® SPSS® Modeler en el modo independiente. Cuando se conecta IBM® SPSS® Modeler a un servidor, utiliza automáticamente la configuración de rendimiento y optimización del servidor.
- Optimizar las rutas para rendimiento máximo. [Si desea obtener más información, consulte el tema Consideraciones de rendimiento para rutas y nodos en el capítulo 13 en Manual de usuario de IBM SPSS Modeler 14.2.](#)

Dispone de más información acerca de cada uno de estos factores del rendimiento en las siguientes secciones.

### ***Configuración de rendimiento y optimización del servidor***

Se pueden configurar ciertos valores de IBM® SPSS® Modeler Server para optimizar el rendimiento. Estos valores se pueden ajustar utilizando la aplicación IBM® SPSS® Collaboration and Deployment Services Deployment Manager, que se incluye con SPSS Modeler Server. [Si desea obtener más información, consulte el tema Uso de IBM SPSS Modeler Administration Console en el capítulo 4 el p. 24.](#)

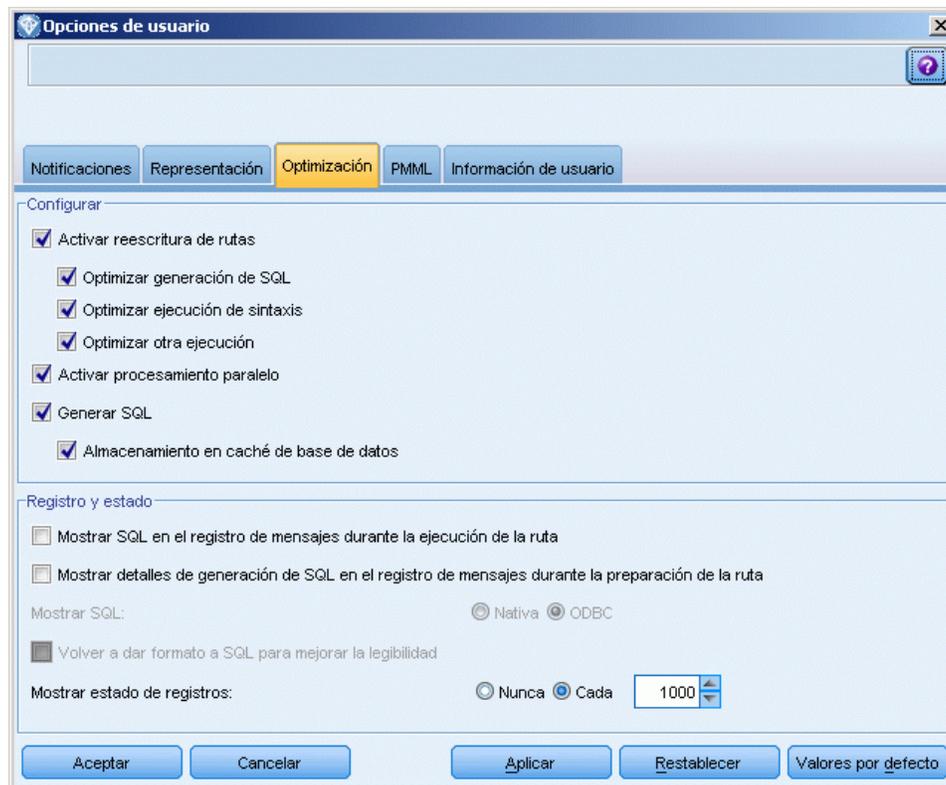
Los valores se agrupan bajo el encabezado Rendimiento y optimización en la ventana de configuración de Deployment Manager. Estos valores están configurados previamente para un rendimiento óptimo en la mayoría de instalaciones. Sin embargo, puede necesitar ajustarlos dependiendo del hardware en concreto, el tamaño de los conjuntos de datos y el contenido de las rutas. [Si desea obtener más información, consulte el tema Rendimiento/Optimización en el capítulo 4 el p. 29.](#)

### ***Configuración de rendimiento y optimización de cliente***

La pestaña Optimización del cuadro de diálogo Opciones de usuario permite optimizar el rendimiento de las rutas. Tenga en cuenta que la configuración de la optimización y el rendimiento en IBM® SPSS® Modeler Server (si se utiliza) reemplaza cualquier configuración en cliente. *Esta pestaña está sólo disponible si se ha activado la conectividad de SPSS Modeler Server en el ordenador IBM® SPSS® Modeler.*

*Nota:* Optimización SQL y modelado de bases de datos requieren que la conectividad de SPSS Modeler Server esté activada en el equipo con SPSS Modeler. Con esta configuración activada, puede acceder a los algoritmos de bases de datos, devolver SQL directamente desde SPSS Modeler y acceder a SPSS Modeler Server. Para verificar el estado de licencia actual, pulse Acerca de en el menú Ayuda de SPSS Modeler. [Si desea obtener más información, consulte el tema Conexión con IBM SPSS Modeler Server en el capítulo 3 en Manual de usuario de IBM SPSS Modeler 14.2.](#)

Figura 5-1  
Cuadro de diálogo Opciones de usuario, pestaña Optimización



*Nota:* La compatibilidad con retrotracción y optimización de SQL depende del tipo de base de datos en uso. Si desea obtener la información más reciente acerca de las bases de datos y los controladores ODBC compatibles y cuyo uso se ha comprobado con IBM® SPSS® Modeler 14.2, vea las matrices de compatibilidad de productos disponibles en el sitio de asistencia corporativo <http://www.ibm.com/support>.

**Activar reescritura de rutas.** Seleccione esta opción para activar la reescritura de rutas en SPSS Modeler. Hay dos tipos de reescritura disponibles, y puede seleccionar (una o ambas). La reescritura reordena los nodos de una ruta en segundo plano para obtener un funcionamiento más eficiente sin alterar la semántica de la ruta.

- **Optimizar generación de SQL.** Esta opción permite reordenar los nodos de la ruta de manera que se puedan retrotraer más operaciones usando la generación de SQL para su ejecución en la base de datos. Si encuentra un nodo que no se puede representar en SQL, el optimizador leerá previamente para comprobar si hay algún nodo por debajo de la ruta que se pueda

representar en SQL y mover con seguridad al nodo problema sin que la semántica de la ruta se vea afectada. No sólo la base de datos puede realizar operaciones de un modo más eficiente que SPSS Modeler, sino que los puntos de retrotracción actúan para reducir el tamaño del conjunto de datos que se devuelve a SPSS Modeler para su procesamiento. A su vez, esto puede reducir el tráfico de red y acelerar las operaciones de las rutas. Tenga en cuenta que la casilla de verificación Generar SQL debe estar seleccionada para que la optimización de SQL tenga algún efecto. [Si desea obtener más información, consulte el tema Optimización de SQL en el capítulo 6 el p. 49.](#)

- **Optimizar ejecución de sintaxis.** Este método de reescritura de rutas incrementa la eficiencia de las operaciones que incorporan más de un nodo con la sintaxis de IBM® SPSS® Statistics. La optimización se logra combinando los comandos de sintaxis en una única operación, en lugar de ejecutar cada operación como separada.
- **Optimizar otra ejecución.** Este método de reescritura de rutas incrementa la eficiencia de las operaciones que no se pueden delegar a la base de datos. La optimización se consigue reduciendo la cantidad de datos de la ruta lo antes posible. Al tiempo que se mantiene la integridad de los datos, la ruta se reescribe para colocar las operaciones más cerca del origen de datos, reduciendo de este modo los datos de la parte de abajo en la ruta en el caso de operaciones costosas, como las uniones.

**Activar procesamiento paralelo.** Cuando trabaje en un ordenador con varios procesadores, esta opción permite al sistema equilibrar la carga a través de estos procesadores, que puede resultar en rendimiento más rápido. El uso de varios nodos, o de los siguientes nodos individuales, puede mejorar gracias al procesamiento paralelo: C5.0, Fundir (por clave), Ordenar, Intervalo (métodos de mosaico y rango), y Agregar (utilizando uno o más campos clave).

**Generar SQL.** Seleccione esta opción para activar la optimización de SQL, permitiendo que las operaciones de ruta se retrotraigan a la base de datos mediante el código SQL para generar los procesos de ejecución, que puede mejorar el rendimiento. Para mejorar aún más el rendimiento, se puede seleccionar Optimizar generación de SQL para maximizar el número de operaciones retrotraídas a la base de datos. Si las operaciones de un nodo se han retrotraído a la base de datos, el nodo se resaltará en color púrpura cuando se ejecute la ruta.

- **Almacenamiento en caché de base de datos.** Para que se ejecuten las rutas que generan SQL en la base de datos, los datos se pueden almacenar en caché en medio de la ruta en una tabla temporal en la base de datos en lugar de en el sistema de archivos. Al combinarlo con la optimización de SQL, se puede mejorar considerablemente el rendimiento. Por ejemplo, el resultado de una ruta que fusiona varias tablas para crear una vista de minería de datos se puede guardar en caché y reutilizar cuando sea necesario. Con el almacenamiento en caché de la base de datos activado, sólo tiene que pulsar con el botón derecho del ratón en cualquier nodo no terminal para almacenar en caché los datos en ese punto, y la caché se creará de forma directa y automática en la base de datos la próxima vez que se ejecute la ruta. Esto permite generar SQL para nodos por debajo de la ruta, mejorando aún más el rendimiento. Si lo prefiere, esta opción se puede desactivar si es necesario, por ejemplo, cuando las directivas o los permisos excluyen los datos que se escriben en la base de datos. Si no se activa el almacenamiento en caché de la base de datos o la optimización de SQL, la caché se escribirá en el sistema de archivos en lugar de en la base de datos. [Si desea obtener más información, consulte el tema Opciones de caché de los nodos en el capítulo 5 en Manual de usuario de IBM SPSS Modeler 14.2.](#)

*Nota:* debido a diferencias menores en la implementación de SQL, las rutas ejecutadas en una base de datos pueden devolver resultados ligeramente diferentes cuando se ejecutan en SPSS Modeler. Por motivos similares, estas diferencias también pueden variar en función del proveedor de la base de datos.

**Mostrar SQL en el registro de mensajes durante la ejecución de la ruta.** Especifica si el SQL generado durante la ejecución de la ruta se debe pasar al registro de mensajes.

**Mostrar detalles de generación de SQL en el registro de mensajes durante la preparación de la ruta.** Durante la presentación preliminar de la ruta, especifica si una presentación preliminar del SQL que se generaría se pasa al registro de mensajes.

**Mostrar SQL.** Especifica si cualquier SQL que aparece en el registro debe contener funciones SQL nativas o funciones ODBC estándar del tipo {fn FUNC(...)} tal como las genera SPSS Modeler. La primera confía en la funcionalidad del controlador ODBC que puede no implementarse. Por ejemplo, este control no tendría efecto alguno para SQL Server.

**Volver a dar formato a SQL para mejorar la legibilidad.** Especifica si se debe dar formato a SQL que se muestre en el registro para mejorar la legibilidad.

**Mostrar estado de registros.** Especifica si los registros se deben generar cuando lleguen a los nodos terminales. Especifique un número que se utilizará para actualizar el estado cada  $N$  registros.

Pulse en Valores por defecto para invertir la configuración del sistema por defecto de esta pestaña.

## ***Optimización y uso de bases de datos***

**Servidor de base de datos.** Si es posible, cree una instancia de base de datos dedicada para la minería de datos, de manera que el servidor de producción no se vea afectado por las consultas de IBM® SPSS® Modeler. Las instrucciones SQL generadas por SPSS Modeler pueden ser muy exigentes: varias tareas del equipo con IBM® SPSS® Modeler Server pueden estar ejecutando SQL en la misma base de datos.

**Minería interna de la base de datos.** Muchos proveedores de bases de datos proporcionan extensiones de minería de datos para sus productos. Tales extensiones permiten que se ejecuten actividades de minería de datos (como generación de modelos o puntuación) en el servidor de la base de datos o en un servidor dedicado diferente. Las funciones de minería interna de base de datos de SPSS Modeler complementan y amplían la capacidad de generación de SQL, proporcionando una manera de controlar las extensiones de la base de datos específicas del proveedor. En algunos casos, este método evita la sobrecarga potencialmente costosa de transferencia de datos entre SPSS Modeler y la base de datos. El almacenamiento en caché de la base de datos puede aumentar aún más las ventajas. Si desea obtener más información, consulte el archivo *DatabaseMiningGuide.pdf*, disponible en la carpeta *\Documentation* que se encuentra en el disco de instalación de IBM® SPSS® Modeler.

## Optimización de SQL

Para obtener el mejor rendimiento, debe intentar siempre maximizar la cantidad de SQL generado para aprovechar el rendimiento y la escalabilidad de la base de datos. Sólo las partes de la ruta que no se pueden compilar en SQL se deben ejecutar en IBM® SPSS® Modeler Server. Si desea obtener más información, consulte [el capítulo 6](#).

### **Carga de datos de archivo**

Los datos que no están almacenados en la base de datos no se pueden beneficiar de la optimización de SQL. Si los datos que desea analizar no están ya en una base de datos, puede cargarlos utilizando un nodo de resultado Base de datos. También se puede utilizar este nodo para almacenar conjuntos de datos intermedios procedentes de la preparación y resultados de implementación de los datos. [Si desea obtener más información, consulte el tema Opciones avanzadas de exportación de base de datos en el capítulo 7 en \*Nodos de origen, proceso y resultado de IBM SPSS Modeler 14.2\*.](#)

IBM® SPSS® Modeler puede interactuar con los cargadores externos para muchos sistemas de base de datos habituales. Se incluye una serie de procesos con el software y están disponibles (con documentación) en el subdirectorio `/scripts` en su carpeta de instalación de SPSS Modeler.

La siguiente tabla muestra las ventajas potenciales de la carga masiva para el rendimiento. La figura muestra el tiempo transcurrido para exportar 250.000 registros y 21 campos a una base de datos de Oracle. El cargador externo era la utilidad `sqlldr` de Oracle.

**Tabla 5-1**  
*Ventajas de la carga masiva para el rendimiento*

Opción de exportación	Tiempo (en segundos)
Por defecto (ODBC)	409
Carga masiva a través de ODBC	52
Carga masiva a través del cargador externo	33

# Optimización de SQL

Una de las características más eficaces de IBM® SPSS® Modeler es la capacidad de realizar multitud de operaciones de preparación y minería de datos directamente en la base de datos. Al generar código SQL que se puede retrotraer a la base de datos para su ejecución, se podrán realizar muchas operaciones (como muestreo, ordenación y derivación de nuevos campos, así como determinados tipos de gráficos) en la base de datos en lugar de en el ordenador de IBM® SPSS® Modeler o IBM® SPSS® Modeler Server. Cuando se trabaja con conjuntos de datos de gran volumen, estos **puntos de retrotracción** pueden mejorar considerablemente el rendimiento de varias maneras:

- Reduciendo el tamaño del conjunto de resultados que se va a transferir de DBMS a SPSS Modeler. Cuando se leen conjuntos de resultados de gran tamaño a través de un controlador ODBC, pueden surgir deficiencias del controlador o de E/S en la red. Por este motivo, las operaciones que más se benefician de la optimización de SQL son la agregación y selección de filas y columnas (nodos Seleccionar, Muestrear, Agregar), que suelen reducir el tamaño del conjunto de datos que se va a transferir. Los datos también pueden almacenarse en la caché de una tabla temporal de la base de datos en puntos críticos de la ruta (por ejemplo, después de un nodo Fundir o Seleccionar) para mejorar aún más el rendimiento.
- Beneficiándose del rendimiento y la escalabilidad de la base de datos. La eficacia se mejora porque un DBMS puede, a menudo, sacar partido del procesamiento paralelo, un hardware más potente, una administración más sofisticada de almacenamiento en disco y la presencia de índices.

Dadas estas ventajas, SPSS Modeler está diseñado para maximizar la cantidad de SQL generado por cada ruta de forma que SPSS Modeler Server sólo ejecuta las operaciones que no se puedan compilar en SQL. No obstante, ciertas operaciones podrían ser incompatibles debido a las limitaciones en lo que se puede expresar en estándar de SQL (SQL 92). [Si desea obtener más información, consulte el tema Sugerencias para maximizar la generación de SQL el p. 55.](#)

*Nota:* Debido a diferencias menores en la implementación de SQL, las rutas ejecutadas en una base de datos pueden devolver resultados ligeramente diferentes cuando se ejecutan en SPSS Modeler. Por motivos similares, estas diferencias también pueden variar dependiendo del proveedor de la base de datos. Por ejemplo, dependiendo de la configuración de la base de datos respecto al uso de mayúsculas y minúsculas al comparar y ordenar cadenas, las rutas de SPSS Modeler ejecutadas mediante retrotracción SQL pueden generar resultados diferentes de los obtenidos al ejecutar dichas rutas sin retrotracción SQL. Póngase en contacto con el administrador de la base de datos para solicitar ayuda sobre la configuración de la base de datos. Para maximizar la compatibilidad con SPSS Modeler, las comparaciones de cadena de la base de datos deben tener en cuenta mayúsculas y minúsculas.

*Nota:* Optimización SQL y modelado de bases de datos requieren que la conectividad de SPSS Modeler Server esté activada en el equipo con SPSS Modeler. Con esta configuración activada, puede acceder a los algoritmos de bases de datos, devolver SQL directamente desde SPSS Modeler y acceder a SPSS Modeler Server. Para verificar el estado de licencia actual, pulse Acerca de en el

menú Ayuda de SPSS Modeler. [Si desea obtener más información, consulte el tema Conexión con IBM SPSS Modeler Server en el capítulo 3 en \*Manual de usuario de IBM SPSS Modeler 14.2\*.](#)

### **Requisitos de base de datos**

Si desea obtener la información más reciente acerca de las bases de datos y los controladores ODBC compatibles y cuyo uso se ha comprobado con SPSS Modeler, vea las matrices de compatibilidad de productos disponibles en el sitio de asistencia corporativo <http://www.ibm.com/support>.

Tenga en cuenta que puede conseguir mejoras de rendimiento adicionales al utilizar el modelado de la base de datos. [Si desea obtener más información, consulte el tema Conceptos básicos del modelado de la base de datos en el capítulo 2 en \*Manual de minería interna de bases de datos de IBM SPSS Modeler 14.2\*.](#)

## **Cómo funciona la generación de SQL**

Los fragmentos iniciales de una ruta de los nodos de origen de la base de datos son los objetivos principales para la generación de SQL. Cuando se encuentra un nodo que no se puede compilar en SQL, los datos se extraen de la base de datos y IBM® SPSS® Modeler Server realiza el procesamiento posterior.

Durante la preparación de una ruta y antes de su ejecución, el proceso de generación de SQL tiene lugar como se indica a continuación:

- El servidor reordena las rutas para mover nodos que se encuentran debajo de la ruta a la “zona de SQL” donde sea seguro realizar esta operación. (Esta función puede desactivarse en el servidor.)
- Si trabaja desde los nodos de origen hacia los nodos terminales, las expresiones SQL se construyen de forma incremental. Esta fase se detiene cuando se detecta que un nodo no se puede convertir a SQL o cuando el nodo terminal (por ejemplo, Tabla o Gráfico) se convierte a SQL. Al final de esta fase, cada nodo se etiqueta con una instrucción SQL si el nodo y sus predecesores tienen un equivalente SQL.
- Si trabaja desde los nodos con los equivalentes de SQL más complejos hacia los nodos de origen, se comprueba la validez de SQL. A continuación se elige el SQL validado correctamente para su ejecución.
- Los nodos para los que todas las operaciones tienen SQL generado se resaltan en púrpura en el lienzo de rutas. En función de los resultados, es posible que desee reorganizar la ruta como considere necesario para aprovechar al máximo la ejecución de la base de datos. [Si desea obtener más información, consulte el tema Sugerencias para maximizar la generación de SQL el p. 55.](#)

### **Dónde se producen mejoras**

La optimización de SQL mejora el rendimiento en determinadas operaciones de datos:

- **Uniones (fundir por clave).** Las operaciones de unión pueden aumentar la optimización de las bases de datos.

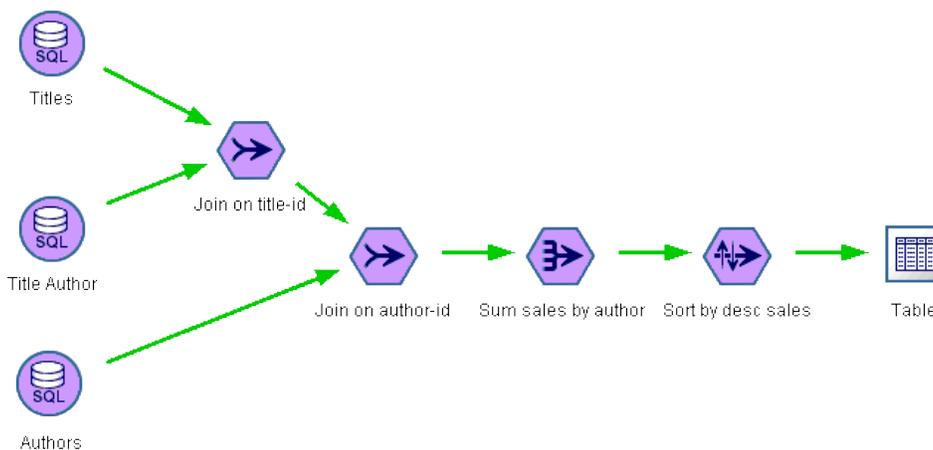
- **Agregación.** Los nodos Agregar, Distribución y Malla utilizan la agregación para generar los resultados. Los datos resumidos utilizan un ancho de banda considerablemente menor al de los datos originales.
- **Selección.** La selección de registros en función de ciertos criterios reduce la cantidad de registros.
- **Ordenación.** La ordenación de los registros es una actividad que requiere un uso intensivo de recursos y que se realiza de una forma más eficaz en una base de datos.
- **Derivación de campos.** Se generan campos nuevos de una forma más eficaz en una base de datos.
- **Proyección de campos.** SPSS Modeler Server extrae sólo los campos necesarios para el siguiente procesamiento de la base de datos, lo que minimiza los requisitos de memoria y ancho de banda. Lo mismo se aplica a los campos superfluos de los archivos planos: aunque el servidor debe leer los datos superfluos, no les asigna almacenamiento.
- **Puntuación.** SQL puede generarse desde árboles de decisión, conjuntos de reglas, regresión lineal y modelos generados por factores.

### Ejemplo de generación de SQL

En la siguiente ruta se unen tres tablas de base de datos por operaciones principales y, a continuación, se realiza una adición y una ordenación.

Figura 6-1

Ruta optimizada con nodos púrpura que indica puntos de retrotracción de SQL (operaciones llevadas a cabo en la base de datos)



### SQL generado

El SQL generado para esta ruta es:

```

SELECT
  T2. au_lname AS C0,
  T2. au_fname AS C1,
  SUM({fn CONVERT(T0. ytd_sales ,SQL_BIGINT)}) AS C2
FROM
  
```

```

dbo . titles T0,
dbo . titleauthor T1,
dbo . authors T2
WHERE
(T0. title_id = T1. title_id )
AND (T1. au_id = T2. au_id )
GROUP BY T2. au_lname ,T2. au_fname
ORDER BY 3 DESC

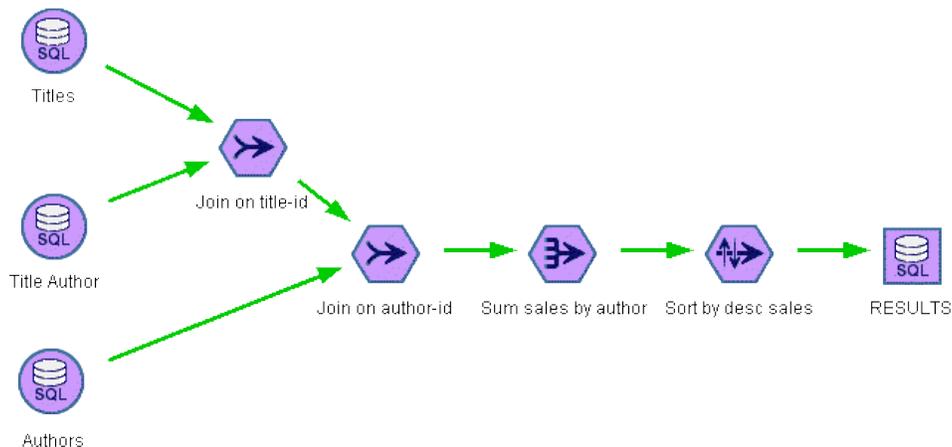
```

### **Ejecución de la ruta**

Cuando se completa la ruta con un nodo de exportación de base de datos, se puede ejecutar la ruta completa en la base de datos.

**Figura 6-2**

*Ruta completa ejecutada en la base de datos*



## **Configuración de la optimización de SQL**

- ▶ Instale un controlador ODBC y configure un origen de datos para la base de datos que desea utilizar. [Si desea obtener más información, consulte el tema Acceso a los datos en el capítulo 2 el p. 11.](#)
- ▶ Cree una ruta que utilice un nodo de origen para obtener datos de dicha base de datos.
- ▶ Compruebe que la generación de SQL está activada en el cliente y el servidor, si es necesario. Está activado por defecto para ambos.

### **Para activar la optimización de SQL en el cliente**

- ▶ En el menú Herramientas, seleccione Opciones de usuario.
- ▶ Pulse en la pestaña Optimización. Seleccione Generar SQL para activar la optimización de SQL. De forma opcional, puede seleccionar otra configuración para mejorar el rendimiento. [Si desea obtener más información, consulte el tema Configuración de rendimiento y optimización de cliente en el capítulo 5 el p. 44.](#)

### **Para activar la optimización de SQL en el servidor**

Dado que la configuración del servidor reemplaza cualquier especificación realizada en el cliente, los valores de configuración del servidor de reescritura de rutas y generación automática de SQL deben estar activados. Si desea obtener más información sobre cómo cambiar la configuración de IBM® SPSS® Modeler Server, consulte la sección [Rendimiento/Optimización en el capítulo 4](#).

### **Para activar la optimización al puntuar modelos**

Para realizar la puntuación, la generación de SQL debe estar activada de forma independiente para cada nodo de modelado, sin tener en cuenta ninguna configuración de servidor o cliente. Esta acción se lleva a cabo porque algunos modelos generan expresiones SQL extremadamente complejas que tal vez no se puedan evaluar eficazmente en la base de datos. Puede que la base de datos indique errores al intentar ejecutar el SQL generado, debido al tamaño o complejidad de SQL.

Puede que se necesite cierto número de tareas de ensayo y error para determinar si la generación de SQL mejora el rendimiento de un modelo determinado. Esta acción se lleva a cabo en la pestaña Configuración después de añadir un modelo generado a una ruta. [Si desea obtener más información, consulte el tema Uso de nugget de modelo en rutas en el capítulo 3 en \*Nodos de modelado de IBM SPSS Modeler 14.2\*](#).

## **Presentación preliminar del SQL generado**

El SQL generado se puede ver en el registro de mensajes antes de su ejecución en la base de datos. Esto puede resultar de utilidad para realizar tareas de depuración y también permite la exportación del SQL generado para su edición o ejecución en la base de datos en un futuro. Además, indica los nodos que se devolverán a la base de datos, lo que puede ayudar a determinar si la ruta se puede volver a ordenar para mejorar el rendimiento.

- ▶ Asegúrese de que las opciones Mostrar SQL en el registro de mensajes durante la ejecución de la ruta y Mostrar detalles de generación de SQL en el registro de mensajes durante la preparación de la ruta están activadas en el cuadro de diálogo Opciones de usuario. [Si desea obtener más información, consulte el tema Configuración de rendimiento y optimización de cliente en el capítulo 5 el p. 44](#).
- ▶ En el lienzo de rutas seleccione el nodo o la ruta cuya presentación preliminar desee ver.
- ▶ Pulse en el botón Presentación preliminar... de la barra de herramientas.

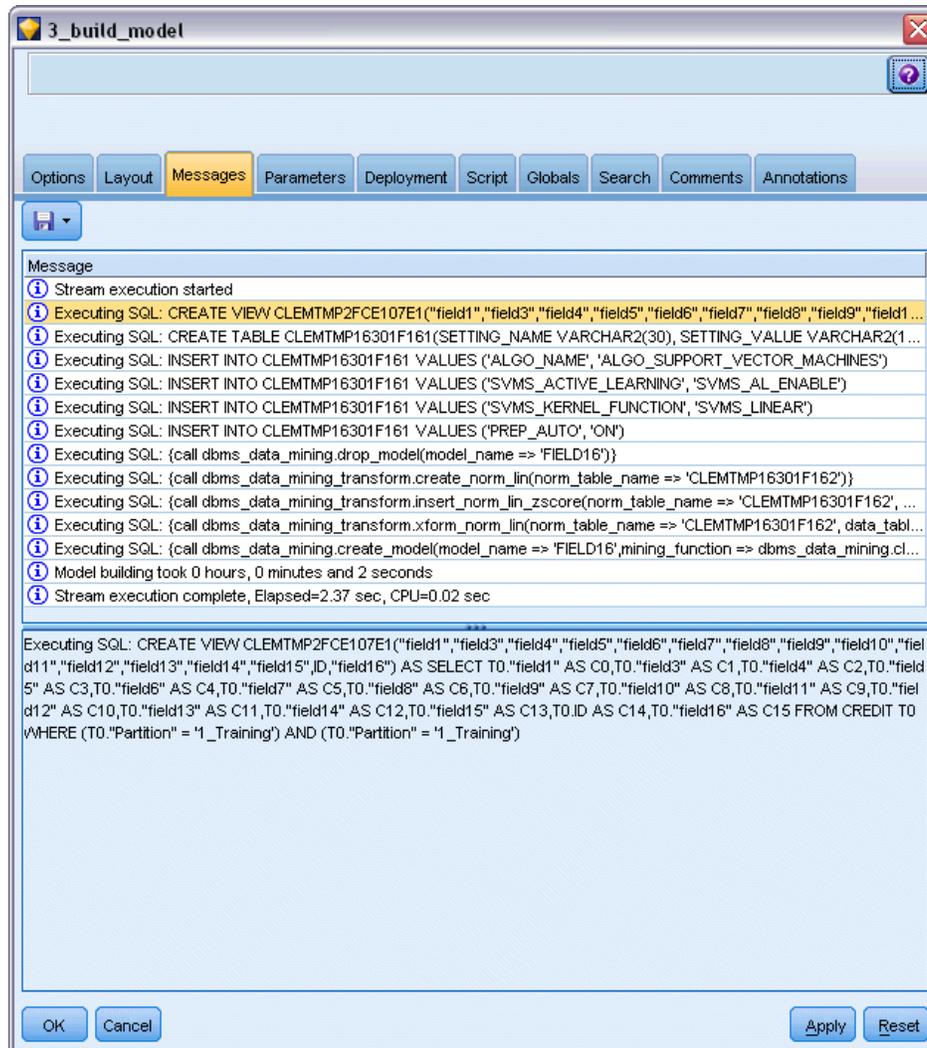
Todos los nodos para los que se genera SQL (y que se devolverán a la base de datos cuando se ejecute la ruta) aparecen en color púrpura en el lienzo de rutas.

Figura 6-3  
Botón de presentación preliminar de SQL



- ▶ Para ver la presentación preliminar del SQL, seleccione en los menús: Herramientas > Propiedades de ruta > Mensajes...

Figura 6-4  
SQL generada como se muestra en el registro de mensajes



## Visualización de SQL de nuggets de modelo

En algunos modelos se puede generar SQL para el nugget de modelo y devuelve la etapa de puntuación de modelo a la base de datos. El uso principal de esta función no es mejorar el rendimiento, sino permitir que las rutas que contienen estos nuggets tengan su SQL devueltos. [Si desea obtener más información, consulte el tema Nodos que admiten la generación de SQL el p. 56.](#)

Para ver la SQL de un nugget de modelo que admite la generación de SQL:

- Seleccione la pestaña Configuración del nugget de modelo.

- ▶ Seleccione una de las opciones Generar con/sin compatibilidad de valores perdidos o Generar SQL para este modelo, como adecuadas.
- ▶ En el menú de nugget de modelo, seleccione:  
File > Exportar SQL
- ▶ Guarde el archivo.
- ▶ Abra el archivo para ver la SQL.

## ***Sugerencias para maximizar la generación de SQL***

Para obtener el mejor aumento del rendimiento de la optimización de SQL, preste atención a los siguientes elementos.

**Orden de rutas.** La generación de SQL puede detenerse cuando la función del nodo no tiene ningún equivalente semántico en SQL debido a que la funcionalidad de minería de datos de IBM® SPSS® Modeler es más importante que las operaciones de procesamiento de datos tradicionales admitidas por SQL estándar. Cuando esto sucede, la generación de SQL también se suprime para cualquier nodo por debajo de la ruta. Por tanto, es posible que pueda mejorar significativamente el rendimiento al reordenar los nodos para colocar las operaciones que detengan SQL en el nivel más inferior de la ruta posible. El optimizador de SQL puede realizar el reordenamiento de forma automática hasta cierto punto (asegúrese de que la reescritura de rutas está activada), pero se pueden realizar más mejoras. Un buen candidato para ello es el nodo Seleccionar, que se puede adelantar. [Si desea obtener más información, consulte el tema Nodos que admiten la generación de SQL el p. 56.](#)

**CLEM expresiones** Si no se puede reordenar una ruta, puede cambiar las opciones de nodo, las expresiones CLEM o, de otro modo, volver a distribuir la forma de realizar la operación, para que no inhiba la generación de SQL. Normalmente, los nodos Derivar, Seleccionar y similares pueden representarse en SQL, siempre y cuando todos los operadores de expresiones CLEM tengan equivalentes en SQL. La mayoría de los operadores se pueden representar, pero hay algunos que inhiben la generación de SQL (en especial, las funciones de secuencia o [“funciones @”]). En ocasiones, la generación se detiene porque la consulta generada se vuelve demasiado compleja para que la base de datos la controle. [Si desea obtener más información, consulte el tema Operadores y expresiones CLEM que admiten la generación de SQL. el p. 60.](#)

**Nodos de varias fuentes.** Cuando una ruta tiene varios nodos de origen de base de datos, la generación de SQL se aplica a cada rama de entrada de manera independiente. Si la generación se detiene en una rama, puede continuar en otra. Cuando dos ramas se fusionan (y ambas pueden expresarse en SQL hasta la fusión), la propia fusión puede a menudo sustituirse por una unión de base de datos y la generación se puede continuar por debajo de la ruta.

**Algoritmos de la base de datos.** La estimación de modelos siempre se lleva a cabo en IBM® SPSS® Modeler Server en lugar de en la base de datos, excepto al utilizar algoritmos nativos de base de datos de Microsoft, IBM u Oracle. [Si desea obtener más información, consulte el tema Conceptos básicos del modelado de la base de datos en el capítulo 2 en \*Manual de minería interna de bases de datos de IBM SPSS Modeler 14.2.\*](#)

**Modelos de puntuación.** La puntuación interna de la base de datos es compatible con algunos modelos al representar el modelo generado en SQL. Sin embargo, algunos modelos generan expresiones SQL extremadamente complejas que no siempre se evalúan de forma eficaz en la base de datos. Por este motivo, la generación de SQL debe estar activada independientemente para cada nodo de modelo. Si descubre que un nodo de modelo está inhibiendo la generación de SQL, vaya a la pestaña Configuración del cuadro de diálogo del nodo y seleccione Generar SQL para este modelo (con algunos modelos, es posible que existan opciones adicionales que controlan la generación). Ejecute comprobaciones para confirmar que la opción es beneficiosa para la aplicación. [Si desea obtener más información, consulte el tema Nodos que admiten la generación de SQL el p. 56.](#)

Al probar nodos de modelado para verificar si la generación de SQL para modelos funciona de manera eficaz, recomendamos que en primer lugar guarde todas las rutas de IBM® SPSS® Modeler. Es posible que algunos sistemas de bases de datos se cuelguen al intentar procesar el SQL generado (potencialmente complejo) y deba cerrar SPSS Modeler desde el Administrador de tareas de Windows.

**Almacenamiento en caché de base de datos.** Si utiliza una caché de nodo para guardar datos en puntos críticos de la ruta (por ejemplo, tras un nodo Fundir o Agregar), asegúrese de que el almacenamiento en caché de base de datos está activado con la optimización de SQL. Esto permite almacenar los datos en caché en una tabla temporal en la base de datos (en lugar de en el sistema de archivos) en la mayoría de los casos. [Si desea obtener más información, consulte el tema Configuración de la optimización de SQL el p. 52.](#) Consulte también [Opciones de caché de los nodos](#) si desea obtener más información.

**SQL específico del proveedor.** La mayor parte del SQL generado se ajusta a los estándares (SQL-92), aunque algunas funciones específicas del proveedor no estándar se explotan en los casos donde son prácticas. El grado de optimización de SQL puede variar en función del origen de base de datos.

## ***Nodos que admiten la generación de SQL***

Las tablas siguientes muestran los nodos que representan operaciones de minería de datos que admiten la generación de SQL. A excepción de los nodos de modelado de la base de datos, si un nodo no aparece en estas tablas, no admite la generación de SQL.

Puede obtener una presentación preliminar de la SQL que se genera antes de ejecutarla. [Si desea obtener más información, consulte el tema Presentación preliminar del SQL generado el p. 53.](#)

### **Fuentes**

<b>Nodo que admite la generación de SQL</b>	<b>Notas</b>
Base de datos	Este nodo se utiliza para especificar tablas y vistas que se pueden usar para realizar análisis adicionales. Este nodo permite la entrada de consultas SQL. Evite conjuntos de resultados con nombres de columna duplicados. <a href="#">Si desea obtener más información, consulte el tema Escritura de consultas SQL el p. 62.</a>

**Operaciones con registros**

Nodo que admite la generación de SQL	Notas
Select	Admite la generación de SQL sólo si se admite también la generación para la propia expresión de selección (consulte a continuación las expresiones). Si algún campo tiene valores nulos, la generación de SQL no ofrece los mismos resultados para descartar, tal y como se proporcionan en IBM® SPSS® Modeler nativo.
Ejemplo	El muestreo simple admite la generación de SQL tal y como se muestra en la tabla siguiente. El muestreo complejo no admite la generación de SQL.
Agregar	
Adición de RFM	Admite la generación, salvo si guarda la fecha de la segunda o tercera transacción más recientes, o sólo si se incluyen transacciones recientes. Sin embargo, la inclusión de las transacciones recientes funciona si se devuelve la función <code>date-time_date(YEAR,MONTH,DAY)</code> / <code>datetime_date(YEAR,MONTH,DAY)</code> .
Sort	
Fundir	No se genera SQL en el caso de fusión por orden. Fusión por clave con unión total o parcial exterior sólo es compatible si el controlador/base de datos lo admite. Puede cambiar el nombre de los campos de entrada no coincidentes mediante un nodo Filtro o la pestaña Filtro de un nodo de origen. En todos los tipos de fusiones, <code>SQL_SP_EXISTS</code> no se admite si las entradas se originan en bases de datos diferentes.
Append	Admite la generación si las entradas no están ordenadas.
Distinguir	

La compatibilidad de la generación de SQL en el nodo Muestrear con muestreo simple se muestra en la tabla siguiente.

Mode	Ejemplo	Tamaños máximo	Semilla	DB2 OS/Z	DB2 OS/400	DB2 Windows / UNIX	Netezza	Oracle	SQL Server	Teradata
Include	First	n/a		Y	Y	Y	Y	Y	Y	Y
	1-de cada-n	off		Y	Y	Y	Y	Y		Y
		max		Y	Y	Y	Y	Y		Y
	Aleatorio %	off	off			Y	Y	Y		Y
			en			Y		Y		
		max	off			Y	Y	Y		Y
		en			Y		Y			
Discard	First	off					Y	Y		
		max					Y	Y		
	1-de cada-n	off		Y	Y	Y	Y	Y		Y
		max		Y	Y	Y	Y	Y		Y
	Aleatorio %	off	off			Y	Y	Y		Y
			en			Y		Y		
		max	off			Y	Y	Y		Y
			en			Y		Y		

**Operaciones con campos**

<b>Nodo que admite la generación de SQL</b>	<b>Notas</b>
Type	Admite la generación de SQL si el nodo Tipo está instanciado y no se especifica ninguna comprobación de tipo ABORTo WARN.
Filter	
Derivar	Admite la generación de SQL si se admite también la generación para la expresión de derivación (consulte a continuación las expresiones).
Conjunto	Admite la generación de SQL de objetivos continuos. Para otros objetivos, admite la generación sólo si se utiliza el método de conjunto "La mayor confianza gana".
Rellenar	Admite la generación de SQL si se admite también la generación para la expresión de derivación (consulte a continuación las expresiones).
Anonimizar	Admite generación de SQL de objetivos continuos y la generación parcial de SQL de objetivos nominales y marca.
Reclasificar	
En intervalo	Admite la generación de SQL si se utiliza el método de intervalos "Cuantiles (igual frecuencia)" y se selecciona la opción "Leer de la pestaña Agrupar valores en intervalos si está disponible".
Análisis de RFM	Admite la generación de SQL si se selecciona la opción "Leer de la pestaña Agrupar valores en intervalos si está disponible", pero los nodos posteriores no lo admitirán.
Partition	Admite la generación de SQL para asignar registros a particiones.
Marcas	
Reestructurar	

**Gráficos**

<b>Nodo que admite la generación de SQL</b>	<b>Notas</b>
Tablero	La generación de SQL es compatible con los siguientes tipos de gráficos: área, áreas en 3D, barras, barras en 3D, barras de recuentos, mapa de calor, sectores, sectores en 3D, sectores de recuentos. En histogramas, la generación de SQL sólo se admite para datos categóricos.
Distribución	
Web	
Evaluación	

**Nodos de modelos generados**

En algunos modelos se puede generar SQL para el nugget de modelo y devuelve la etapa de puntuación de modelo a la base de datos. El uso principal de esta función no es mejorar el rendimiento, sino permitir que las rutas que contienen estos nuggets tengan su SQL devueltos. Si

desear obtener más información, consulte el tema [Visualización de SQL de nuggets de modelo](#) el p. 54.

<b>Nugget de modelo compatible con la generación de SQL</b>	<b>Notas</b>
Árbol C&R	Admite la generación SQL para la opción de árbol único, pero no para las opciones de aumento, agregación autodocimante y conjuntos de datos de gran tamaño.
QUEST	
CHAID	
C5.0	
Lista de decisiones	
Lineal	Admite la generación de SQL para la opción de modelo estándar, pero no para las opciones de aumento, agregación autodocimante y conjuntos de datos de gran tamaño.
Red neuronal	Admite la generación de SQL para la opción de modelo estándar (sólo perceptrón multicapa), pero no para las opciones de aumento, agregación autodocimante y conjuntos de datos de gran tamaño.
PCA/Factorial	
Logística	Admite la generación de SQL para el procedimiento multinomial, pero no binomial. En el procedimiento multinomial, no se admite la generación si se seleccionan las confianzas, salvo que el tipo de objetivo sea Marca.
Conjuntos de reglas generados	

### **Resultado**

<b>Nodo que admite la generación de SQL</b>	<b>Notas</b>
Tabla	Admite la generación si SQL se admite para la expresión de resultado (consulte a continuación las expresiones).
Matriz	Admite la generación excepto si se selecciona “Todos los numéricos” en la opción Campos.
Análisis	Admite la generación en función de las opciones seleccionadas.
Transformar	
Estadísticos	Admite la generación si no se usa la opción Correlacionar.
Informe	
Val. globales	

### **Exportar**

<b>Nodo que admite la generación de SQL</b>	<b>Notas</b>
Base de datos	
Publisher	La ruta publicada contendrá el SQL generado.

## **Operadores y expresiones CLEM que admiten la generación de SQL.**

Estas tablas muestran las operaciones y expresiones matemáticas que admiten la generación de SQL y que a veces se usan durante la minería de datos. Los operadores ausentes en estas tablas no admiten la generación de SQL en la versión actual.

### **Operadores**

<b>Operación que admite la generación de SQL</b>	<b>Notas</b>
+	
-	
/	
*	
><	Se usa para concatenar cadenas.

### **Operadores relacionales**

<b>Operación que admite la generación de SQL</b>	<b>Notas</b>
=	
/=	Se usa para especificar “no es igual”.
>	
>=	
<	
<=	

### **Funciones**

<b>Operación que admite la generación de SQL</b>	<b>Notas</b>
abs	
allbutfirst	
allbutlast	
and	
arccos	
arcsin	
arctan	
arctanh	
cos	
div	
exp	
fracof	
hasstartstring	
hassubstring	
integer	
intof	

Operación que admite la generación de SQL	Notas
isalphacode	
islowercode	
isnumbercode	
isstartstring	
issubstring	
isuppercode	
last	
length	
locchar	
log	
log10	
lowertoupper	
max	
member	
min	
negate	
not	
number	
or	
pi	
real	
rem	
round	
sign	
sin	
sqrt	
string	
strmember	
subscrs	
substring	
substring_between	
uppertolower	
to_string	

### ***Funciones especiales***

Operación que admite la generación de SQL	Notas
@NULL	
@GLOBAL_AVE	Las funciones especiales globales se usan para recuperar valores globales calculados por el nodo Val. globales.
@GLOBAL_SUM	
@GLOBAL_MAX	
@GLOBAL_MEAN	

Operación que admite la generación de SQL	Notas
@GLOBAL_MIN	
@GLOBALSDEV	

### **Funciones de agregación**

Operación que admite la generación de SQL	Notas
Sum	
Mean	
Min	
Max	
Count	
SDev	

### **Uso de funciones SQL en expresiones CLEM**

El @SQLFN se puede utilizar para añadir funciones SQL con nombre en expresiones de CLEM, sólo para su ejecución en bases de datos. Puede ser de gran utilidad en casos especiales en los que es necesario un SQL patentado o cualquier otra personalización específica del proveedor.

El uso de esta función no queda cubierto por el acuerdo de soporte estándar de IBM® SPSS® Modeler, ya que la ejecución depende de componentes de la base de datos externa que escapan al control de IBM Corp., pero puede distribuirse en casos especiales, normalmente como parte de un contrato de servicios. Si es necesario, póngase en contacto con <http://www.ibm.com/software/analytics/spss/services/> para obtener más información.

### **Escritura de consultas SQL**

Cuando se usa el nodo Base de datos, se debe prestar especial atención a cualquier consulta SQL que resulte en un conjunto de datos con nombres de columna duplicados. Estos nombres duplicados evitan a menudo la optimización de SQL para cualquier nodo de la parte de abajo de la ruta.

IBM® SPSS® Modeler usa instrucciones SELECT anidadas para retrotraer SQL en rutas que usan una consulta SQL en el nodo de origen Base de datos. En otras palabras, la ruta anida la consulta especificada en el nodo de origen Base de datos dentro de una o varias instrucciones SELECT generadas durante la optimización de nodos hacia abajo en la ruta. Por tanto, si el conjunto de resultados de una consulta contiene nombres de columna duplicados, la instrucción no puede ser anidada por RDBMS. Las dificultades de anidación surgen más a menudo durante una unión de tablas cuando se selecciona una columna con el mismo nombre en varias de las tablas unidas. Por ejemplo, observe esta consulta en el nodo de origen:

```
SELECT e.ID, e.LAST_NAME, d.*
FROM EMP e RIGHT OUTER JOIN
DEPT d ON e.ID = d.ID;
```

La consulta evitará cualquier optimización de SQL posterior, puesto que esta instrucción **SELECT** resultará en un conjunto de datos con dos columnas denominadas **ID**.

Para permitir la optimización completa de SQL, se debe ser más explícito al escribir consultas SQL y especificar alias de columna cuando surja una situación de conflicto por nombres de columna duplicados. La siguiente instrucción ejemplifica una consulta más explícita:

```
SELECT e.ID AS ID1, e.LAST_NAME, d.*  
FROM EMP e RIGHT OUTER JOIN  
DEPT d ON e.ID = d.ID;
```

# ***Configuración de Oracle para plataformas UNIX***

## ***Configuración de Oracle para la optimización de SQL***

Cuando se ejecuta IBM® SPSS® Modeler Server en plataformas UNIX y que leen la información de una base de datos de Oracle, tenga en cuenta las siguientes sugerencias para garantizar que el SQL generado se optimiza totalmente dentro de la base de datos.

### ***Especificación de configuración regional apropiada***

Cuando ejecute SPSS Modeler Server en una ubicación regional distinta a la de procedencia de los controladores de Connect ODBC, debe volver a configurar el equipo para mejorar la optimización de SQL. Los controladores de Connect ODBC se envían sólo con los archivos de configuración regional *en\_US* (inglés de Estados Unidos). Por lo tanto, si el equipo de SPSS Modeler Server se está ejecutando con una configuración regional diferente o si el shell en que se inició SPSS Modeler Server no tiene definida totalmente la configuración regional, puede que el SQL generado no esté totalmente optimizado en Oracle. Los motivos son los siguientes:

- SPSS Modeler Server usa los archivos de configuración regional de ODBC correspondientes a la configuración regional en la que se ejecuta para traducir a cadenas de texto los códigos devueltos desde la base de datos. A continuación utiliza estas cadenas de texto para determinar a qué base de datos se está conectando actualmente.
- Si la configuración regional (tal como devuelve a SPSS Modeler Server la consulta \$LANG del sistema) no es *en\_US*, IBM® SPSS® Modeler no puede traducir a texto los códigos que recibe del controlador ODBC. En otras palabras, un código no traducido en lugar de la cadena *Oracle* se devuelve a SPSS Modeler Server al comienzo de una conexión base de datos. Esto quiere decir que SPSS Modeler no puede optimizar las rutas por Oracle.

Para comprobar y restablecer las especificaciones de configuración regional:

- ▶ En un shell de UNIX, ejecute:

```
#locale
```

Se devolverá información de configuración regional para el shell. Por ejemplo:

```
$ locale
LANG=en_US.ISO8859-15
LC_CTYPE="en_US.ISO8859-15"
LC_NUMERIC="en_US.ISO8859-15"
LC_TIME="en_US.ISO8859-15"
LC_COLLATE="en_US.ISO8859-15"
LC_MONETARY="en_US.ISO8859-15"
LC_MESSAGES="en_US.ISO8859-15"
```

```
LC_ALL=en_US.ISO8859-15
```

- ▶ Cambie a su directorio de configuración regional/Connect ODBC. (Aquí verá un único directorio: *en\_US*).
- ▶ Cree un enlace flexible a este directorio *en\_US*, especificando el nombre de la configuración regional en el shell. A continuación se muestra un ejemplo:

```
#ln -s en_US en_US.ISO8859-15
```

En el caso de configuraciones locales distintas al inglés, como *fr\_FR.ISO8859-1*, debe crear el enlace flexible de esta forma:

```
#ln -s en_US fr_FR.ISO8859-1
```

- ▶ Una vez que haya creado el vínculo, reinicie SPSS Modeler Server desde este mismo shell. (SPSS Modeler Server recibe la información de configuración regional del shell desde el que se inició.)

### **Notas**

Al optimizar una máquina UNIX para puntos de retrotracción de SQL a Oracle, tenga en cuenta estas sugerencias:

- Se debe especificar totalmente la configuración regional. En el ejemplo anterior, debe crear el enlace en el formato *idioma\_territorio.código-página*. El directorio de configuración regional existente *en\_US* no es suficiente.
- Para optimizar totalmente la minería interna de la base datos, tanto *LANG* como *LC\_ALL* se deben definir en el shell que se use para iniciar SPSS Modeler Server. *LANG* se puede definir en el shell como se haría con cualquier otra variable de entorno antes de reiniciar SPSS Modeler Server. Por ejemplo, consulte la siguiente definición:

```
#LANG=en_US.ISO8859-15; export LANG
```

- Cada vez que inicie SPSS Modeler Server, necesitará comprobar que la información de configuración regional del shell está completamente especificada y que existe el enlace flexible correspondiente en el directorio de configuración regional/ODBC.

# Configuración de procesos de inicio de UNIX

## Introducción

En este apéndice se describen algunos de los procesos que se incluyen con las versiones UNIX de IBM® SPSS® Modeler Server y, asimismo, se explica el modo de configurarlos. Los procesos se utilizan para:

- Configurar SPSS Modeler Server para iniciarse automáticamente cuando el equipo servidor se reinicie.
- Detener automáticamente SPSS Modeler Server y reiniciarlo.
- Cambiar permisos de los archivos creados con SPSS Modeler Server.
- Configurar SPSS Modeler Server para que funcione con los controladores de Connect ODBC proporcionados con SPSS Modeler Server. [Si desea obtener más información, consulte el tema IBM SPSS Modeler Server y el paquete Data Access Pack el p. 68.](#)
- Configurar SPSS Modeler Server para utilizar IBM® SPSS® Collaboration and Deployment Services Enterprise View Driver al leer datos con un nodo Enterprise View. [Si desea obtener más información, consulte el tema Configuración de un controlador para el nodo Enterprise View el p. 75.](#)

## Procesos

IBM® SPSS® Modeler Server utiliza varios procesos, entre otros:

- **modelersrv.sh.** El proceso de inicio manual de SPSS Modeler Server se encuentra en el directorio de instalación de SPSS Modeler Server. Configura el entorno del servidor cuando el proceso demonio del servidor se inicia *manualmente*. Ejecútelo cuando desee iniciar y detener el servidor de forma manual. Edítelo cuando necesite cambiar la configuración para iniciar el sistema manualmente.
- **auto.sh.** Se trata de un proceso mediante el que se configura el sistema para iniciar automáticamente el proceso demonio del servidor durante el arranque. Ejecútelo una sola vez para configurar el sistema para que se inicie automáticamente. No es necesario editarlo. El proceso se encuentra en el directorio de instalación de SPSS Modeler Server.
- **rc.modeler.** Cuando *auto.sh* se ejecuta, se crea este proceso en una ubicación que depende del sistema operativo del servidor. Configura el entorno del servidor cuando éste se inicia *automáticamente*. Edítelo cuando necesite cambiar la configuración para iniciar el sistema automáticamente.

Tabla B-1  
Ubicación de *rc.modeler* según el sistema operativo

Sistema operativo	Ubicación
AIX	<i>/etc/rc.modeler</i>
HP-UX	<i>/sbin/init.d/rc.modeler</i>
Solaris	<i>/etc/init.d/rc.modeler</i>

## Inicio y parada automáticos de IBM SPSS Modeler Server

IBM® SPSS® Modeler Server debe iniciarse como un proceso demonio. El programa de instalación incluye un proceso (*auto.sh*) que puede ejecutar para configurar el sistema para que detenga y reinicie SPSS Modeler Server de manera automática.

### Para configurar el sistema para iniciarse y apagarse automáticamente

- ▶ Inicie sesión como root.
- ▶ Cambie al directorio de instalación de SPSS Modeler Server.
- ▶ Ejecute el proceso. En el símbolo del sistema de UNIX, escriba:

```
./auto.sh
```

Se creará un proceso de inicio automático (*rc.modeler*) en la ubicación que indica la tabla anterior. El sistema operativo usará *rc.modeler* para iniciar el proceso demonio de SPSS Modeler Server cuando el equipo servidor se reinicie. El sistema operativo también utilizará *rc.modeler* para detener el centinela cuando el sistema esté apagado.

## Inicio y parada manuales de IBM SPSS Modeler Server

Puede iniciar y detener manualmente IBM® SPSS® Modeler Server ejecutando el proceso *modelersrv.sh*.

### Para iniciar y detener IBM SPSS Modeler Server manualmente

- ▶ Cambie al directorio de instalación de SPSS Modeler Server.
- ▶ Para iniciar el servidor, en el símbolo del sistema de UNIX, escriba:

```
./modelersrv.sh start
```

- ▶ Para detener el servidor, en el símbolo del sistema de UNIX, escriba:

```
./modelersrv.sh stop
```

## Edición de procesos

Si emplea el inicio tanto manual como automático, realice los mismos cambios en *modelersrv.sh* y en *rc.modeler*. Si solamente utiliza el inicio manual, efectúe los cambios en *modelersrv.sh*. Si sólo utiliza el inicio automático, realice los cambios en *rc.modeler*.

**Para editar los procesos**

- ▶ Detenga IBM® SPSS® Modeler Server. (Si desea obtener más información, consulte el tema [Inicio y parada manuales de IBM SPSS Modeler Server el p. 67.](#))
- ▶ Busque el proceso adecuado. (Si desea obtener más información, consulte el tema [Procesos el p. 66.](#))
- ▶ Abra el proceso en un editor de texto, realice los cambios y guarde el archivo.
- ▶ Inicie SPSS Modeler Server, ya sea automáticamente (reiniciando el equipo servidor) o de forma manual.

**Control de permisos en la creación de archivos**

IBM® SPSS® Modeler Server crea archivos temporales con permisos de lectura, escritura y ejecución para todos. Puede omitir este comportamiento por defecto si edita el valor `UMASK` en el proceso de inicio, ya sea en `modelersrv.sh`, `rc.modeler` o en ambos. (Si desea obtener más información, consulte [Edición de procesos](#) anteriormente.) Recomendamos usar `077` como el valor `UMASK` más restrictivo. Aquellos valores que sean más restrictivos podrían provocar problemas relacionados con los permisos de SPSS Modeler Server.

**IBM SPSS Modeler Server y el paquete Data Access Pack**

Si desea usar los controladores de Connect for ODBC con IBM® SPSS® Modeler Server, el entorno de Connect for ODBC debe estar configurado mediante `odbc.sh` cuando se inicie el proceso de SPSS Modeler Server. Para ello, puede editar el proceso de inicio de IBM® SPSS® Modeler pertinente, ya sea en `modelersrv.sh`, en `rc.modeler` o en ambos. (Si desea obtener más información, consulte el tema [Edición de procesos el p. 67.](#))

Si desea obtener más información, consulte la página Web corporativa en <http://www.spss.com/drivers/>. Si tiene alguna pregunta acerca de la creación o configuración de permisos de los orígenes de datos ODBC, póngase en contacto con el administrador de la base de datos.

**Para configurar Connect for ODBC para iniciarse con IBM SPSS Modeler Server**

- ▶ Desde el sitio Web corporativo en <http://www.spss.com/drivers/clientCLEM.htm>, descargue el archivo tar comprimido correspondiente de la plataforma en la que ha instalado SPSS Modeler Server. Asegúrese de descargar los controladores para la versión correcta de SPSS Modeler Server que ha instalado. Copie el archivo a la ubicación en la que desea instalar los controladores ODBC (por ejemplo, `/usr/spss/odbc`).
- ▶ Extraiga el archivo.
- ▶ Ejecute el proceso `setodbcpath.sh` que se ha extraído del archivo.

- Modifique el proceso *odbc.sh* para agregar la definición de ODBCINI a la parte inferior de este proceso y expórtelo, por ejemplo:

```
ODBCINI=/usr/spss/odbc/odbc.ini; exportar ODBCINI
```

ODBCINI debe apuntar al nombre de ruta completo del archivo *odbc.ini* que desea que lea SPSS Modeler para obtener una lista de los orígenes de datos de ODBC que defina (se ha instalado un *odbc.ini* por defecto con los controladores).

- Guarde *odbc.sh*.
- (Sólo para la instalación de SPSS Modeler Server de 64 bits; para otras instalaciones, continúe desde el paso siguiente) Defina y exporte LD\_LIBRARY\_PATH\_64 en *odbc.sh*:

```
if [ "$LD_LIBRARY_PATH_64" = "" ]; then
    LD_LIBRARY_PATH_64=<library_path>
else
    LD_LIBRARY_PATH_64=<library_path>:$LD_LIBRARY_PATH_64
fi
export LD_LIBRARY_PATH_64
```

donde *library\_path* es lo mismo que para la definición de LD\_LIBRARY\_PATH del proceso que se ha inicializado con su ruta de instalación (por ejemplo, */usr/spss/odbc/lib*). La manera más fácil de hacerlo es copiar las instrucciones *if* y *export* para LD\_LIBRARY\_PATH en su archivo *odbc.sh*, adjuntarlas al final del archivo y, a continuación, sustituir las cadenas "LD\_LIBRARY\_PATH" en las últimas instrucciones *if* y *export* que se han añadido con "LD\_LIBRARY\_PATH\_64".

Así, su archivo *odbc.sh* final en una instalación de SPSS Modeler Server de 64 bits puede tener la siguiente apariencia:

```
—if [ "$LD_LIBRARY_PATH" = "" ]; then
    LD_LIBRARY_PATH=/usr/spss/odbc/lib
else
    LD_LIBRARY_PATH=/usr/spss/odbc/lib:$LD_LIBRARY_PATH
fi
export LD_LIBRARY_PATH
if [ "$LD_LIBRARY_PATH_64" = "" ]; then
    LD_LIBRARY_PATH_64=/usr/spss/odbc/lib
else
    LD_LIBRARY_PATH_64=/usr/spss/odbc/lib:$LD_LIBRARY_PATH_64
fi
export LD_LIBRARY_PATH_64
ODBCINI=/usr/spss/odbc/odbc.ini; export ODBCINI
```

No olvide exportar LD\_LIBRARY\_PATH\_64, así como definirlo con el bucle *if*.

- Modifique el archivo *odbc.ini* que ha definido anteriormente con \$ODBCINI. Defina los nombres de origen de datos que necesite (estos nombres dependen de la base de datos a la que esté accediendo).
- Guarde el archivo *odbc.ini*.

- ▶ Configure SPSS Modeler Server para utilizar estos controladores. Para ello, modifique *modelersrv.sh* y añada la siguiente línea inmediatamente a continuación de la línea que define SCLEMDNAME:

```
. <odbc.sh_path>
```

donde *odbc.sh\_path* es la ruta completa al archivo *odbc.sh* que ha modificado al principio de este procedimiento, por ejemplo:

```
./usr/spss/odbc/odbc.sh
```

*Nota:* Aquí la sintaxis es importante; asegúrese de dejar un espacio entre el primer periodo y la ruta al archivo.

- ▶ Guarde *modelersrv.sh*.

#### **Para probar la conexión**

- ▶ Reinicie SPSS Modeler Server.
- ▶ Conéctese a SPSS Modeler Server desde un cliente.
- ▶ En el cliente, agregue un nodo de origen Base de datos al lienzo.
- ▶ Abra el nodo y verifique que puede ver los nombres de origen de datos que ha definido en el archivo *odbc.ini* anterior en el procedimiento de configuración.

Si no ve aquí lo que esperaba, u obtiene errores al intentar conectarse a un origen de datos que ha definido, siga el procedimiento de solución de problemas. [Si desea obtener más información, consulte el tema Solución de problemas de la configuración de Connect for ODBC el p. 71.](#)

#### **Para configurar Connect for ODBC para iniciarse con IBM SPSS Modeler Solution Publisher Runtime**

Cuando pueda conectarse correctamente a la base de datos desde SPSS Modeler Server, puede configurar una instalación de IBM® SPSS® Modeler Solution Publisher Runtime en el mismo servidor haciendo referencia al mismo proceso de *odbc.sh* desde el proceso de inicio de SPSS Modeler Solution Publisher Runtime.

- ▶ Modifique el proceso *modelerrun* en SPSS Modeler Solution Publisher Runtime para añadir la siguiente línea inmediatamente por encima de la última línea del proceso:

```
. <odbc.sh_path>
```

donde *odbc.sh\_path* es la ruta completa al archivo *odbc.sh* que ha modificado al principio de este procedimiento, por ejemplo:

```
./usr/spss/odbc/odbc.sh
```

*Nota:* Aquí la sintaxis es importante; asegúrese de dejar un espacio entre el primer periodo y la ruta al archivo.

- ▶ Guarde el archivo del proceso *modelerrun*.

**Para configurar Connect for ODBC para iniciarse con IBM SPSS Modeler Batch**

No es necesario configurar el proceso de IBM® SPSS® Modeler Batch para ODBC. Se debe a que se conecta a SPSS Modeler Server desde SPSS Modeler Batch para ejecutar rutas. Asegúrese de que se ha realizado la configuración ODBC de SPSS Modeler Server y de que está funcionando correctamente, tal y como se ha descrito anteriormente.

**Para añadir o modificar un nombre de origen de datos**

- ▶ Modifique el archivo *odbc.ini* para incluir el nombre nuevo o modificado.
- ▶ Pruebe la conexión tal y como se ha descrito anteriormente.

Si la conexión con SPSS Modeler Server funciona correctamente, el origen de datos nuevo o modificado debería funcionar también correctamente con SPSS Modeler Solution Publisher Runtime y SPSS Modeler Batch.

**Solución de problemas de la configuración de Connect for ODBC****No se muestran orígenes de datos ni texto aleatorio**

Si abre un nodo de origen Base de datos y la lista de orígenes de datos disponibles está vacía o contiene entradas inesperadas, es posible que se deba a un problema con el proceso de inicio.

- ▶ Compruebe que se haya definido \$ODBCINI en *modelersrv.sh*, tanto explícitamente en el proceso en sí, como en el proceso *odbc.sh* al que se hace referencia en *modelersrv.sh*.
- ▶ En el caso siguiente, asegúrese de que ODBCINI apunta a la ruta completa del archivo *odbc.ini* que ha utilizado para definir sus orígenes de datos de ODBC.
- ▶ Si la especificación de ruta en ODBCINI es correcta, compruebe el valor de \$ODBCINI que se está utilizando en el entorno de IBM® SPSS® Modeler Server mostrando la variable en *modelersrv.sh*. Para ello, añada la línea siguiente a *modelersrv.sh* después del punto en el que defina ODBCINI:

```
mostrar $ODBCINI
```

- ▶ Guarde y ejecute *modelersrv.sh* a continuación. El valor de \$ODBCINI que se está estableciendo en el entorno de SPSS Modeler Server se escribe en *stdout* para su verificación.
- ▶ Si no se devuelve ningún valor a *stdout* y si está definiendo \$ODBCINI en el proceso de *odbc.sh* al que hace referencia desde *modelersrv.sh*, compruebe que la sintaxis de referencia sea la correcta. Debería ser:

```
.<odbc.sh_path>
```

donde *odbc.sh\_path* es la ruta completa al archivo *odbc.sh* que ha modificado al principio de este procedimiento, por ejemplo:

```
./usr/spss/odbc/odbc.sh
```

*Nota:* Aquí la sintaxis es importante; asegúrese de dejar un espacio entre el primer periodo y la ruta al archivo.

Cuando se muestra el valor correcto a *stdout* en la ejecución de *modelersrv.sh*, debería poder ver los nombres de origen de datos en el nodo de origen Base de datos cuando reinicia SPSS Modeler Server y se conecta a él desde el cliente.

***El cliente de IBM SPSS Modeler se cuelga al pulsar en el cuadro de diálogo Connect en Conexiones a la base de datos***

Este comportamiento puede deberse a que su ruta de la biblioteca que no se ha definido correctamente para incluir la ruta a las bibliotecas de ODBC. La ruta de la biblioteca se define por `$LD_LIBRARY_PATH` (y `$LD_LIBRARY_PATH_64` en versiones de 64 bits).

Para ver el valor de la ruta de la biblioteca en el entorno de daemon de SPSS Modeler Server, muestre el valor de la variable de entorno adecuada desde *modelersrv.sh* después de la línea donde está agregando la ruta de la biblioteca de ODBC a la ruta de la biblioteca, y ejecute el proceso. Se mostrará el valor de la ruta de la biblioteca al terminal cuando ejecute a continuación el proceso.

Si está haciendo referencia a *odbc.sh* desde *modelersrv.sh* para configurar su entorno de ODBC de SPSS Modeler Server, muestre el valor de la ruta de su biblioteca desde la línea posterior desde donde hace referencia al proceso de *odbc.sh*. Para mostrar el valor, añada la siguiente línea al proceso, y guarde y ejecute el archivo del proceso:

```
echo $<variable_ruta_biblioteca>
```

donde *<variable\_ruta\_biblioteca>* es la variable de la ruta de la biblioteca adecuada para el sistema operativo de su servidor.

El valor devuelto de su ruta de biblioteca debe incluir la ruta al subdirectorio *lib* de su instalación de ODBC. Si no es así, añada esta ubicación al archivo.

Si está ejecutando la versión de 64 bits de SPSS Modeler Server, `$LD_LIBRARY_PATH_64` ignorará `$LD_LIBRARY_PATH`, si se ha establecido. Si está teniendo este problema en una de estas plataformas de 64 bits, muestre `LD_LIBRARY_PATH_64` así como `$LD_LIBRARY_PATH` desde *modelersrv.sh* y, si es necesario, defina `$LD_LIBRARY_PATH_64` de modo que incluya el subdirectorio *lib* de su instalación de ODBC, y exporte la definición.

***No se ha encontrado el nombre de origen de datos y no se ha especificado ningún controlador por defecto***

Si ve este error al pulsar en el cuadro de diálogo Connect en Conexiones a la base de datos, indicará normalmente que su archivo *odbc.ini* se ha definido de manera incorrecta. Compruebe que el nombre de origen de datos (DSN) tal y como se define en la sección [ODBC Data Sources] al principio del archivo, coincide con la cadena especificada entre corchetes más abajo en *odbc.ini* para definir el DSN. Si difieren de alguna manera, verá este error cuando intente conectarse utilizando el DSN en IBM® SPSS® Modeler. A continuación se muestra un ejemplo de una especificación *incorrecta*:

```
[Orígenes de datos ODBC]  
Oracle=Oracle Wire Protocol
```

```
....
....
[Controlador de Oracle]
Driver=/usr/ODBC/lib/XEora22.so
Description=SPSS 5.2 Oracle Wire Protocol
AlternateServers=
....
```

Debe cambiar una de las dos cadenas en negrita para que la coincidencia sea exacta. De esta manera se solucionará el error.

### **No se ha podido cargar el controlador especificado**

Este error también indica que el archivo *odbc.ini* se ha definido de manera incorrecta. Una posibilidad es que el parámetro Controlador de la estrofa del controlador se haya definido de manera incorrecta, por ejemplo:

```
[Orígenes de datos ODBC]
Oracle=Oracle Wire Protocol

....
....
[Oracle]
Driver=/nosuchpath/ODBC/lib/XEora22.so
Description=SPSS 5.2 Oracle Wire Protocol
AlternateServers=
```

- ▶ Compruebe que el objeto compartido especificado por el parámetro Controlador existe.
- ▶ Corrija la ruta del objeto compartido si es incorrecta.
- ▶ Si se especifica el parámetro Controlador en este formato:

```
Driver=ODBCHOME/lib/XEora22.so
```

indicará que no ha inicializado sus procesos relacionados con ODBC. Ejecute el proceso *setodbcpath.sh* que se ha instalado con los controladores. [Si desea obtener más información, consulte el tema IBM SPSS Modeler Server y el paquete Data Access Pack el p. 68.](#) Cuando haya ejecutado este proceso, debería ver que la cadena “ODBCHOME” se ha sustituido por la ruta que lleva a su instalación de ODBC. Esto debería resolver el problema.

Otra causa puede deberse a un problema con la biblioteca del controlador. Use la herramienta *ivtestlib* que se proporciona con Connect for ODBC para confirmar que el controlador no se puede cargar. Para Connect64, utilice la herramienta *ddtestlib*. Solucione el problema configurando la variable de la ruta de biblioteca en el proceso de inicio.

Por ejemplo, si no puede cargarse el controlador de Oracle para una instalación de 32 bits, siga estos pasos:

- ▶ Utilice *ivtestlib* para confirmar que el controlador no se puede cargar. Por ejemplo, en el símbolo del sistema de UNIX, escriba:

```
sh
```

```
cd ODBCDir
. odbcs.sh
./bin/ivtestlib MFor815
```

donde ODBCDir se reemplaza con la ruta del directorio de instalación de Connect for ODBC.

- ▶ Lea el mensaje para ver si existe algún error. Un mensaje como el siguiente, por ejemplo:  
Load of MFor815.so failed: ld.so.1: bin/ivtestlib: fatal: libclntsh.so: open failed: No such file or directory indica que falta la biblioteca cliente de Oracle, *libclntsh.so*, o bien que no se encuentra en la ruta de la biblioteca (por ejemplo, no en LD\_LIBRARY\_PATH en Solaris).
- ▶ Confirme que la biblioteca existe. Si no es así, vuelva a instalar el cliente de Oracle. Si la biblioteca está ahí, escriba la siguiente secuencia de comandos desde el símbolo del sistema de UNIX:

```
LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/bigdisk/oracle/product/8.1.6/lib
export LD_LIBRARY_PATH
./bin/ivtestlib Mfor815
```

donde */bigdisk/oracle/product/8.1.6/lib* se reemplaza con la ruta a *libclntsh.so* y LD\_LIBRARY\_PATH es la variable de la ruta de biblioteca del sistema operativo.

Tenga en cuenta que si está ejecutando SPSS Modeler de 64 bits en Linux o Solaris, la variable de la ruta de biblioteca contendrá el sufijo *\_64*. Por lo tanto, las dos primeras líneas del ejemplo anterior serán:

```
LD_LIBRARY_PATH_64=$LD_LIBRARY_PATH_64:/bigdisk/oracle/product/8.1.6/lib
export LD_LIBRARY_PATH_64
```

- ▶ Lea el mensaje para confirmar que ya se puede cargar el controlador. Un mensaje como el siguiente, por ejemplo:  
Load of MFor815.so successful, qehandle is 0xFF3A1BE4 indica que la biblioteca cliente de Oracle se puede cargar.
- ▶ Corrija la ruta de la biblioteca en el proceso de inicio de SPSS Modeler.
- ▶ Reinicie SPSS Modeler Server con el proceso de inicio que ha editado (*modelersrv.sh* o *rc.modeler*).

## Rutas de bibliotecas

El nombre de la variable de la ruta de biblioteca varía en función del sistema operativo. La siguiente tabla se incluye a modo de guía para ayudarle a realizar las sustituciones pertinentes al configurar el sistema o solucionar problemas.

Tabla B-2  
Ruta de biblioteca por sistema operativo

Sistema operativo	Nombre de variable de ruta de biblioteca
AIX	LIBPATH
HP-UX	SHLIB_PATH
Solaris o Linux de 32 bits	LD_LIBRARY_PATH
Solaris o Linux de 64 bits	LD_LIBRARY_PATH_64

## **Configuración de un controlador para el nodo Enterprise View**

El nodo Enterprise View permite crear y mantener una conexión entre una sesión de IBM® SPSS® Modeler y IBM SPSS Collaboration and Deployment Services Enterprise View en un IBM® SPSS® Collaboration and Deployment Services Repository compartido. [Si desea obtener más información, consulte el tema Almacenamiento y recuperación de objetos de IBM SPSS Collaboration and Deployment Services Repository en el capítulo 9 en \*Manual de usuario de IBM SPSS Modeler 14.2\*.](#)

Para utilizar el nodo Enterprise View con IBM® SPSS® Modeler Server, deberá instalar y configurar un IBM® SPSS® Collaboration and Deployment Services Enterprise View Driver. Si desea obtener más información sobre la instalación del controlador, póngase en contacto con el administrador local. Tras la instalación, debe añadirse una referencia a *pev.sh* al proceso de inicio—ya sea en *modelersrv.sh*, *rc.modeler* o en ambos—con el siguiente formato:

```
./usr/odbc/pev.sh
```



# Ejecución como un proceso no root en UNIX.

## Introducción

Estas instrucciones ofrecen información sobre la ejecución de IBM® SPSS® Modeler Server como un proceso no root en sistemas UNIX.

**Ejecución como root.** La instalación por defecto de SPSS Modeler Server supone que el proceso demonio del servidor se ejecutará como root. La ejecución como root permite a IBM® SPSS® Modeler autenticar fiablemente cada inicio de sesión de un usuario e iniciar cada sesión de un usuario en la cuenta de usuario de UNIX correspondiente, lo que garantiza que los usuarios sólo tengan acceso a sus propios archivos y directorios. Sin embargo, es posible ejecutar el proceso demonio en una cuenta no root siguiendo las instrucciones que aparecen a continuación.

**Ejecución como no root.** Sin privilegio root, SPSS Modeler autentica usuarios utilizando una base de datos privada de contraseñas, diferente de la base de datos de contraseñas de UNIX, y comienza todas las sesiones de los usuarios en la misma cuenta de usuario de UNIX. Por tanto, cualquier dato del archivo escrito o leído por SPSS Modeler lo comparten todos los usuarios de SPSS Modeler. El acceso a los datos de la base de datos no se ve afectado ya que los usuarios tienen que autenticarse a sí mismos independientemente en cada origen de datos de la base de datos que utilicen.

## Configuración de IBM SPSS Modeler Server como un proceso no root.

La ejecución de IBM® SPSS® Modeler Server como un proceso no root significa tener los identificadores de usuario efectivo y real del proceso demonio del servidor establecidos en una cuenta de su elección. Aún así, se debe iniciar sesión como root, o su de root, para iniciar o detener el servidor y realizar la administración general.

Si desea configurar SPSS Modeler Server para que se ejecute en una cuenta que no sea root, siga estos pasos:

1. Cree un grupo que contenga todos los usuarios. Puede llamar a este grupo como desee, pero en este ejemplo, lo llamaremos *usuariosmodeler*.
2. Cree una cuenta de usuario en la que ejecutar SPSS Modeler Server. Esta cuenta es para uso único del proceso demonio de SPSS Modeler Server. Es decir, no puede utilizar esta cuenta para iniciar y detener el servidor o para realizar administración general de IBM® SPSS® Modeler. En este ejemplo, le llamaremos *modelerserv*.

Al crear la cuenta, tenga en cuenta que:

- La cuenta no necesita la capacidad de iniciar sesión.

- El grupo principal debería ser el grupo `<usuariosmodeler>>` creado anteriormente.
  - El directorio principal puede ser el directorio de instalación de SPSS Modeler o cualquier otro por defecto conveniente (considere utilizar un directorio diferente al de instalación si necesita que la cuenta perdure a través de las actualizaciones).
3. A continuación, configure los procesos de inicio para iniciar SPSS Modeler Server utilizando la cuenta que acaba de crear. Busque el proceso de inicio adecuado y ábralo en un editor de texto. [Si desea obtener más información, consulte el tema Procesos en el apéndice B el p. 66.](#)

- ▶ En el proceso apropiado, cambie la línea que establece la variable `MODELERUSER` para hacer referencia al nombre de la cuenta de usuario creada para ejecutar SPSS Modeler Server. Por ejemplo:

```
MODELERUSER=modelerserv
```

- ▶ Cambie la configuración de `umask` para permitir al menos acceso de lectura del grupo a los archivos creados:

```
umask 027
```

4. Edite el archivo de opciones del servidor, `config/options.cfg`, para especificar la autenticación respecto a la base de datos privada de contraseñas, añadiendo la línea:

```
authentication_methods, "pasw_modeler"
```

5. A continuación, tendrá que crear una base de datos privada de contraseñas almacenada en el archivo `config/passwords.cfg`. El archivo de contraseñas define las combinaciones de nombre/contraseña de usuario que pueden iniciar sesión en SPSS Modeler. *Nota:* éstas son privadas para SPSS Modeler y no tienen conexión con los nombres y contraseñas de usuarios utilizadas para iniciar sesión en UNIX. Se pueden utilizar los mismos nombres de usuarios por comodidad, pero no se pueden utilizar las mismas contraseñas.

Para crear el archivo de contraseña, necesitará utilizar el programa de utilidad de contraseña, `pwutil`, que se encuentra en el directorio `bin` de la instalación de SPSS Modeler Server. La sinopsis de este programa es:

```
pwutil [ nombre de usuario [ contraseña ] ]
```

El programa toma un nombre de usuario y una contraseña en texto plano y escribe el nombre de usuario y la contraseña cifrada para el resultado estándar en un formato apropiado para su inclusión en el archivo de contraseña. Por ejemplo, para definir un usuario `modeler` con la contraseña “minería de datos” escribiría:

```
bin/pwutil modeler "minería de datos" > config/passwords.cfg
```

En la mayoría de los casos, si todos los usuarios inician sesión con el mismo nombre y la misma contraseña, basta con definir un único nombre de usuario. Sin embargo, se pueden crear usuarios adicionales utilizando el operador `>>` para añadir cada uno al archivo, por ejemplo:

```
bin/pwutil modeler "minería de datos2" >> config/passwords.cfg
```

*Nota:* si se utiliza un único >, se sobrescribirá cada vez el contenido de *passwords.cfg*, reemplazando cualquier usuario establecido anteriormente. Recuerde que todos los usuarios comparten la misma cuenta de usuario de UNIX independientemente.

6. Cambie recursivamente el propietario del directorio de instalación de SPSS Modeler y su contenido para que sea el usuario *<modelerserv>* y grupo *<usuariosmodeler>* donde los nombres a los que se hace referencia son aquellos creados anteriormente. Por ejemplo:

```
chown -R -h modelerserv:usuariosmodeler.
```

7. Considere crear subdirectorios en el directorio de datos para los usuarios de SPSS Modeler de manera que tengan un sitio donde almacenar los datos de trabajo sin interferencias. Estos directorios deberían ser propiedad del grupo *<usuariosmodeler>* y tener permisos de grupo de búsqueda, escritura y lectura. Por ejemplo, para crear un directorio de trabajo para el usuario *bob*:

```
mkdir data/bob
chown bob:modelerserv data/bob
chmod ug=rwx,o= data/bob
```

Adicionalmente, puede establecer el bit establecer-ID-grupo del directorio de manera que cualquier archivo de datos copiado en el directorio pase a ser automáticamente propiedad del grupo *<usuariosmodeler>*:

```
chmod g+s data/bob
```

### ***Ejecución de IBM SPSS Modeler Server como un proceso no raíz en AIX***

Al iniciar SPSS Modeler Server como un proceso no raíz en AIX, existe la posibilidad de que el error *no se encuentra lsattr* se imprimirá en la consola. Se debe a que *lsattr* se encuentra en el directorio */usr/sbin*, pero si SPSS Modeler Server registra un nuevo usuario (para que se ejecute como no raíz) restablece la ruta para incluir únicamente */usr/bin* y */bin*.

Si se produce este error, la solución consiste en crear el siguiente enlace en */usr/bin*:

```
su
cd /usr/bin
ln -s /usr/sbin/lsattr lsattr
```

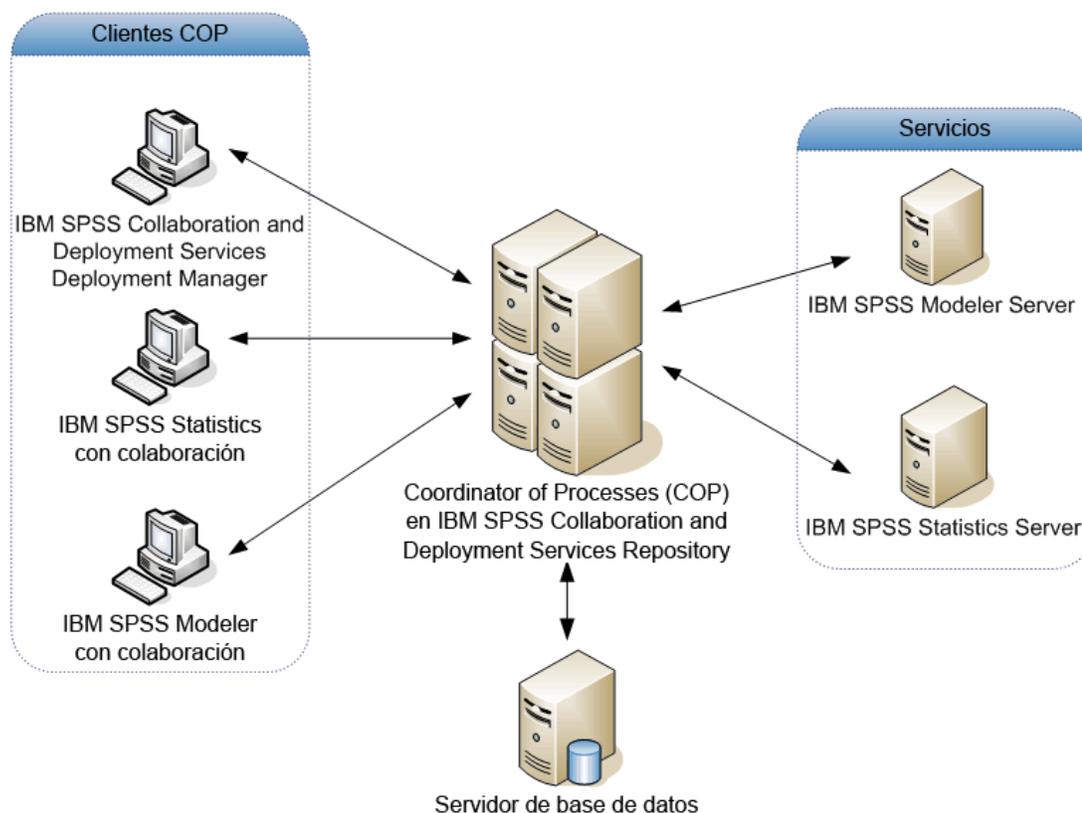
Después de crear el enlace, reinicie SPSS Modeler Server.

## Equilibrado de cargas con conglomerados de servidores

Con IBM® SPSS® Collaboration and Deployment Services, un complemento denominado Coordinator of Processes que puede utilizarse para gestionar servicios en la red. Coordinator of Processes ofrece funciones de administración de servidores diseñadas para optimizar la comunicación y el procesamiento entre cliente y servidor.

Los servicios que se van a gestionar, como IBM® SPSS® Statistics Server o IBM® SPSS® Modeler Server, se registran con Coordinator of Processes en el inicio y envían mensajes de estado actualizados periódicamente. Los servicios pueden almacenar cualquier archivo de configuración necesario en el IBM® SPSS® Collaboration and Deployment Services Repository y recuperarlos cuando se inicia.

Figura D-1  
Arquitectura de Coordinator of Processes



Si ejecuta las rutas de IBM® SPSS® Modeler en un servidor puede aumentar el rendimiento. En algunos casos, sólo puede elegir entre uno o dos servidores. En otros casos, puede ofrecerle una mayor selección de servidores porque hay una diferencia significativa entre cada servidor,

como propietarios, derechos de acceso, datos de servidor, comprobación frente a servidores de producción, etc. Además, si dispone de Coordinator of Processes en su red, es posible que se le ofrezca un conglomerado de servidores.

Un conglomerado de servidores es un grupo de servidores intercambiables en términos de configuración y recursos. Coordinator of Processes determina qué servidor se adapta mejor para responder a una solicitud de procesamiento, mediante un algoritmo que equilibrará la carga según varios criterios, entre los que se incluyen las ponderaciones de servidor, las prioridades de usuario y las cargas de procesamiento actual. Si desea obtener más información, consulte *Coordinator of Processes Service Developer's Guide* disponible bajo pedido en IBM Corp.

Si se conecta a un servidor o conglomerado de servidores en SPSS Modeler, puede introducir un servidor de forma manual o buscar un servidor o conglomerado mediante Coordinator of Processes. [Si desea obtener más información, consulte el tema Conexión con IBM SPSS Modeler Server en el capítulo 3 en \*Manual de usuario de IBM SPSS Modeler 14.2\*.](#)

# Avisos

This information was developed for products and services offered worldwide.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

*IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785, U.S.A.*

For license inquiries regarding double-byte character set (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

*Intellectual Property Licensing, Legal and Intellectual Property Law, IBM Japan Ltd., 1623-14, Shimotsuruma, Yamato-shi, Kanagawa 242-8502 Japan.*

**El párrafo siguiente no se aplica en el Reino Unido ni en cualquier otro país en los que dichas provisiones sean contrarias a la legislación local:** SPSS INC., AN IBM COMPANY, PROPORCIONA ESTA PUBLICACIÓN “TAL CUAL” SIN GARANTÍAS DE NINGÚN TIPO, YA SEA EXPRESAS O IMPLÍCITAS, INCLUYENDO, SIN LIMITAR LA GENERALIDAD DE LAS GARANTÍAS IMPLÍCITAS DE NO INFRACCIÓN, COMERCIALIZACIÓN O IDONEIDAD PARA UN FIN DETERMINADO. Algunos estados no permiten el descargo de responsabilidad de garantías expresas o implícitas en determinadas transacciones, por lo que esta declaración no será aplicable.

Esta información puede incluir imprecisiones técnicas o errores tipográficos. La información que se contiene se puede modificar periódicamente; estos cambios se incorporarán en las nuevas ediciones de la publicación. SPSS Inc. puede realizar mejoras y/o cambios en el producto(s) y/o el programa(s) descrito en esta publicación en cualquier momento sin notificación.

Las referencias a esta información en sitios web ajenos a SPSS y a IBM se proporcionan únicamente por motivos de comodidad y no servirán de ninguna forma como aprobación de esos sitios web. Los materiales de esos sitios web no forman parte de los materiales de este producto de SPSS Inc. y el uso de esos sitios web se realiza bajo su responsabilidad.

Al enviar información a IBM o SPSS, el usuario concede a IBM y a SPSS el derecho no exclusivo de utilizar o distribuir la información de la forma que estime adecuada sin incurrir en obligaciones con el usuario.

La información relacionada con productos ajenos a productos SPSS se ha obtenido de los proveedores de esos productos, de sus anuncios publicados u otros orígenes disponibles de forma pública. SPSS no ha comprobado esos productos y no puede confirmar la precisión del rendimiento, compatibilidad o cualquier otras reclamaciones relacionadas con productos ajenos a SPSS. Las cuestiones sobre las responsabilidades de productos ajenos a SPSS se deben dirigir a los proveedores de esos productos.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

*IBM Software Group, Attention: Licensing, 233 S. Wacker Dr., Chicago, IL 60606, USA.*

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Esta información contiene ejemplos de datos e informes utilizados en operaciones comerciales habituales. Para ilustrarlas de la forma más completa posible, los ejemplos incluyen los nombres de personas, empresas, marcas y productos. Todos estos nombres son inventados y cualquier similitud con los nombres y direcciones de una empresa real es una coincidencia.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

### ***Marcas comerciales***

IBM, el logotipo de IBM e [ibm.com](http://www.ibm.com) son marcas comerciales de IBM Corporation, registradas en múltiples jurisdicciones en todo el mundo. Existe una lista actualizada de las marcas comerciales de IBM disponible en Internet en <http://www.ibm.com/legal/copytrade.shtml>.

SPSS es una marca comercial de SPSS Inc., an IBM Company, registradas en múltiples jurisdicciones en todo el mundo.

---

Adobe, el logotipo de Adobe, PostScript y el logotipo de PostScript son marcas comerciales registradas o marcas comerciales de Adobe Systems Incorporated en los Estados Unidos y/o en otros países.

IT Infrastructure Library es una marca comercial registrada de la Agencia central de telecomunicaciones y computación central que ahora forma parte de la Oficina de comercio gubernamental.

Intel, el logotipo de Intel logo, Intel Inside, el logotipo de Intel, Intel Centrino, el logotipo de Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium y Pentium son marcas comerciales o marcas comerciales registradas de Intel Corporation o de sus filiales en los Estados Unidos y en otros países.

Linux es una marca comercial registrada de Linus Torvalds en los Estados Unidos, en otros países o ambos.

Microsoft, Windows, Windows NT y el logotipo de Windows son marcas comerciales de Microsoft Corporation en los Estados Unidos, en otros países o ambos.

ITIL es una marca comercial registrada y una marca comercial comunitaria registrada de la Oficina de Comercio Gubernamental y está registrada en la Oficina de patentes y marcas comerciales de los Estados Unidos.

UNIX es una marca comercial registrada de The Open Group en los Estados Unidos y en otros países.

Cell Broadband Engine es una marca comercial de Sony Computer Entertainment, Inc. en los Estados Unidos, en otros países o ambos y se utiliza con licencia.

Java y todas las marcas comerciales y logotipos basados en Java son marcas comerciales de Sun Microsystems, Inc. en los Estados Unidos, en otros países o ambos.

Linear Tape-Open, LTO, the LTO Logo, Ultrium, and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and other countries.

Otros nombres de productos y servicios pueden ser marcas comerciales de IBM, SPSS u otras empresas.

---

# Índice

- acceso a datos, 11
- acceso del administrador
  - de IBM SPSS Modeler Server, 25
- adición de conexiones de IBM SPSS Modeler Server, 16–17
- administración
  - de IBM SPSS Modeler Server, 23
  - de ubicación de licencia de IBM SPSS Statistics, 36
- administración de memoria
  - opciones de administración, 29
- allow\_modelling\_memory\_override
  - archivo *options.cfg*, 29
- almacenamiento en caché de base de datos
  - controlling from options.cfg, 34
  - generación de SQL, 56
- almacenamiento en caché del nodo
  - escritura en base de datos, 56
  - generación de SQL, 56
- almacenamiento en caché en la base de datos, 34
- almacenamiento en caché, en la base de datos, 34
- archivo *messages.log*, 42
- archivo *options.cfg*, 34
- archivos de datos
  - IBM SPSS Statistics, 13
  - importación y exportación, 13
- archivos de registro
  - de IBM SPSS Modeler Server, 27, 42
  - presentación del SQL generado, 47, 53
- archivos temporales, 9
  - permisos para (IBM SPSS Modeler Server), 19
- arquitectura
  - componentes, 6
- autenticación, 19
- autenticación de usuario, 19
- auto.sh* (UNIX)
  - ubicación de, 66
- avisos legales, 81
  
- bases de datos
  - acceso, 11
- búsqueda de conexiones en COP, 17
  
- cache\_compression
  - archivo *options.cfg*, 29
- chemsrv.sh* (UNIX)
  - ubicación de, 66
- cifrado
  - SSL, 36
- cifrado de datos de SSL
  - activación de IBM SPSS Modeler Server, 31
- compresión de caché, 29
- conexiones
  - al IBM SPSS Modeler Server, 14, 16–17
  - conglomerado de servidores, 17
  - conexiones a la base de datos
    - cierre, 35
  - configuración de coordinator of processes
    - de IBM SPSS Modeler Server, 32
  - configuración de COP
    - de IBM SPSS Modeler Server, 32
  - configuración del cortafuegos
    - archivo *options.cfg*, 30
  - configuración del puerto
    - archivo *options.cfg*, 29
  - configuración del puerto del servidor
    - archivo *options.cfg*, 29
  - Connect for ODBC
    - configuración en UNIX, 68
  - Coordinator of Processes, 17
  - conglomerados de servidores, 79
    - equilibrado de cargas, 79
  - COP, 17
    - conglomerados de servidores, 79
    - equilibrado de cargas, 79
  - cop\_enabled
    - archivo *options.cfg*, 32
  - cop\_host
    - archivo *options.cfg*, 32
  - cop\_password
    - archivo *options.cfg*, 32
  - cop\_port\_number
    - archivo *options.cfg*, 32
  - cop\_service\_description
    - archivo *options.cfg*, 32
  - cop\_service\_host
    - archivo *options.cfg*, 32
  - cop\_service\_name
    - archivo *options.cfg*, 32
  - cop\_service\_weight
    - archivo *options.cfg*, 32
  - cop\_update\_interval
    - archivo *options.cfg*, 32
  - cop\_user\_name
    - archivo *options.cfg*, 32
- cuentas de usuario
  - IBM SPSS Modeler Server, 19
  
- data\_file\_path
  - archivo *options.cfg*, 28
- data\_files\_restricted
  - archivo *options.cfg*, 28
- database
  - support tiers, 11
- DB2
  - optimización de SQL, 49–50
- detención de IBM SPSS Modeler Server
  - en UNIX, 21
  - en Windows, 21

- directorio temporal
  - de IBM SPSS Modeler Server, 28
- discos duros, 10
- documentación, 2
- ejecución de varias rutas, 30
- ejemplos
  - conceptos básicos, 4
  - Manual de aplicaciones, 2
- ejemplos de aplicaciones, 2
- error en la ejecución de la ruta, 30
- espacio en disco, 9
  - cálculo, 10
- estado
  - de IBM SPSS Modeler Server en UNIX, 21
  - de IBM SPSS Modeler Server en Windows, 21
- expresiones CLEM
  - generación de SQL, 60
- Función @SQLFN, 62
- generación de SQL, 44, 49–50, 56
  - activación, 52
  - activación de IBM SPSS Modeler Server, 31
  - expresiones CLEM, 55, 60
  - presentación preliminar, 47, 53
  - reescritura de rutas, 55
  - registro, 47, 53
  - sugerencias, 55
  - visualización de nuggets de modelo, 54
- IBM SPSS Modeler, 1
  - documentación, 2
- IBM SPSS Modeler Administration Console, 24
  - acceso del administrador, 25
- IBM SPSS Modeler Server
  - acceso del administrador, 25
  - administración de, 24
  - archivos de registro, 27
  - autenticación de usuario, 19
  - configuración de coordinator of processes, 32
  - configuración de COP, 32
  - creación de archivos, 19
  - cuentas de usuario, 19
  - directorio temporal, 28
  - ID de usuario, 14
  - información para usuarios finales, 18
  - nombre de dominio (Windows), 14
  - nombre de host, 14, 16
  - número de puerto, 14, 16, 27
  - opciones de administración, 23
  - opciones de configuración, 27
  - password, 14
  - procesos del servidor, 33
  - procesos que no responden, 22
  - resultados diferentes de Client, 20
  - supervisión del uso, 33
- IBM SPSS Statistics archivos de datos
  - importación y exportación, 13
- IBM SPSS Text Analytics, 2
- ID de usuario
  - IBM SPSS Modeler Server, 14
- inicio automático del servidor
  - configuración en UNIX, 67
- inicio de IBM SPSS Modeler Server
  - en UNIX, 21
  - en Windows, 21
- inicio de sesión en IBM SPSS Modeler Server, 14
- inicio único de sesión, 16
- io\_buffer\_size
  - archivo *options.cfg*, 29
- LDAP, 41
  - seguridad, 41
- límites de kernel de UNIX, 22
- límites de kernel en UNIX, 22
- log\_database\_access
  - archivo *options.cfg*, 27
- log\_database\_errors
  - archivo *options.cfg*, 27
- log\_directory
  - archivo *options.cfg*, 27
- log\_file\_access
  - archivo *options.cfg*, 27
- log\_file\_errors
  - archivo *options.cfg*, 27
- marcas comerciales, 82
- max\_file\_size
  - archivo *options.cfg*, 28
- max\_login\_attempts
  - archivo *options.cfg*, 27
- max\_parallelism
  - archivo *options.cfg*, 29
- max\_sessions
  - archivo *options.cfg*, 27
- max\_sql\_string\_length
  - archivo *options.cfg*, 31
- memoria, 10
- memory\_usage
  - archivo *options.cfg*, 29
- mensajes
  - presentación del SQL generado, 47, 53
- Microsoft SQL Server
  - optimización de SQL, 49–50
- minería interna de la base de datos, 47
- modelado
  - administración de memoria, 29
- modelling\_memory\_limit\_percentage
  - archivo *options.cfg*, 29



- 
- nombres de columna duplicados, 62
  - optimización de Oracle, 64
  - presentación generada, 53
  - visualización de nuggets de modelo, 54
  - SQL Server
    - optimización de SQL, 49–50
  - sql\_generation\_enabled
    - archivo *options.cfg*, 31
  - SSL, 36
    - comunicaciones seguras, 37
    - conceptos básicos, 36
  - ssl\_certificate\_file
    - archivo *options.cfg*, 31
  - ssl\_enabled
    - archivo *options.cfg*, 31
  - ssl\_private\_key\_file
    - archivo *options.cfg*, 31
  - ssl\_private\_key\_password
    - archivo *options.cfg*, 31
  - stream\_rewriting\_enabled
    - archivo *options.cfg*, 29
  
  - Tecnología de acceso a datos de IBM SPSS Statistics, 11
  - temp\_directory
    - archivo *options.cfg*, 28
  - tiers, database support, 11
  
  - ubicación de licencia de IBM SPSS Statistics, 36
  - UNIX
    - autenticación de usuario, 19
    - configuración de permisos de archivo, 68
    - rutas de bibliotecas, 74
  
  - vista previa
    - generación de SQL, 53